

Chirac's newly won winning ways

In the aftermath of the French election the consequences for science of the demise of the ministry for research may be overestimated.

THE new Prime Minister of France, M. Jacques Chirac, is inevitably less demonstrative a supporter of the French research and development establishment than was his predecessor, M. Laurent Fabius, who had succeeded to that post from his earlier appointment as minister of research. So much was widely predicted in the weeks preceding the election, when M. Chirac's advisers were forever saying that the research bureaucracy had become a natural target for their cost-cutting decentralizing zeal (*Nature* 319, 729; 1986). Even so, it will have come as a shock to many people that these widely flagged intentions could have implied anything so radical as the abolition of the ministry of research (see page 295). Henceforth, unless M. Chirac changes tack, research will be no better distinguished from the other routine matters with which governments are concerned than it is, for example, in Britain. Is this not a sign that the new government intends for science the policies of malign neglect on which previous governments in Paris have habitually relied? And will not the result be that the resurgence of the past five years will be more rapidly put into reverse?

Fears such as these will be widely expressed in the next few weeks, but it is too soon to take them seriously. M. Chirac may intend to keep his election promise to let the French ministry of research go into limbo. It remains to be seen whether he will have the courage (some would say gall) also to reorganize along the lines of his manifesto the agencies by which French research has been supported over the past five or even fifty years, the research councils of which the Centre National de la Recherche Scientifique (CNRS) is the chief. Before the election, M. Chirac's advisers were promising that the agricultural and medical research councils would be transferred to the corresponding ministries and that CNRS, by far the largest, would be broken into smaller pieces. To some extent, the motivation of these policies is doctrinal, born of the conviction more common on the right than the left that governments may create the framework for industrial innovation but can have no substantial (or useful) part to play in the execution of industrial strategy. Yet there is no suggestion that the French agricultural and health research councils will be much transformed by what M. Chirac plans for them. The crucial issues are not administrative but financial.

Tradition

The future of CNRS is bound to be more difficult to decide. Traditionally, even before the Second World War, CNRS had several distinct but parallel functions. In recent decades, CNRS has become best known as the agency by which major efforts in basic science are undertaken, in fields as different as high-energy physics and geochemistry. Innovations of this kind have spawned the distinctive French efforts in space research and oceanography, now administratively independent. But CNRS is also the traditional partner, with the universities, in the conduct of academically based research, but in a manner that is interestingly distinct from that practised almost everywhere else. Most of its work in support of university research is channelled through the full-time CNRS employees who work alongside academics in university laboratories, exciting both the admiration of colleagues and their envy. On the face of things, it would

not be a disaster if some of this activity were transformed into that more familiar elsewhere in Western Europe and North America, the provision of research funds to university researchers by means of the familiar competition for research grants. Siting the administration of CNRS within the ministry of education in Paris will make it easier for M. Chirac and his colleagues to contemplate a change along these lines. Provided that the contemplation is serious, not a token prelude to predetermined and prejudiced reorganization, there may even be benefits to be won from change in this direction. Robbing CNRS of its role as the chief sponsor of major projects in basic research would be a more dangerous step to contemplate at this early stage.

Change

The other obvious consequences of the post-election reorganization of the French government are inevitably harder to assess, if only because of the novelty (and ambiguity) of the arrangement that has harnessed a socialist president to a right-wing executive in the constitution of the government. President Mitterrand will be anxious to ensure that the executive and the French Assembly do not chip away at his constitutional authority over French foreign policy, which will make the whole area of French collaboration with other countries a potential no-man's land in the months ahead. The chances are that the President will be able to sustain those French policies stemming from his personal enthusiasm for European collaboration in high technology. It would not be surprising if the Prime Minister, as a *quid pro quo*, were to succeed in aligning France with Britain in scepticism about the need for continued investment in collaborative high-energy physics through the agency of CERN, the high-energy physics laboratory at Geneva. Provided that both partners in this uneasy government alliance acknowledge that the immense benefits that have accrued to France in the past five years are a prize that cannot be lightly thrown away, even developments such as these need not spell doom.

Newly elected governments tend to brim over with enthusiasm better suited to routine matters of public administration than to the proper care for sensitive parts of public life, among which research may be the most delicate. It is one thing to nationalize (or denationalize) a string of banks or other businesses, replacing one bunch of shareholders by another, and quite another to pretend to know what decisions should be made, and how, in a field in which even the practitioners are at a loss to tell just what needs doing. So what M. Chirac must keep clearly in mind, during the first few heady weeks when enthusiasm may get the better of his colleagues' judgement, is that France has done extremely well in basic research during the past decade for reasons which are not sufficiently explained by the way in which money has been thrown at some researchers during the past five years. During the same period, for reasons which are probably quite separate, the French telecommunications network has ceased to be a joke and become something of a marvel instead. Nobody's interest would be served if the new government, by pretending that it knows not merely the questions but the answers, were to make a mess of this beneficent development. □



SDI

Japan eager to join in the star wars fray

Tokyo

JAPAN is sailing steadily towards taking part in the US Strategic Defense Initiative (SDI), and the opposition that was expected to wreck participation has altogether failed to materialize. At the end of this week, 46 representatives and engineers from 21 of Japan's high-technology companies, accompanied by nine government officials, will set off to the United States for a 10-day tour of institutes and companies where SDI research is taking place. This is the first time that industry has been involved in SDI missions to the United States, and completes the groundwork for industrial participation.

Executives of high-technology companies are already happy to admit that they have been contacted by government officials encouraging them to look for SDI contracts. What is worrying industry now is not whether they will be allowed to participate but, as the head of one small electronics company put it, "the extent to which the government is intending to try to control affairs from behind the scenes".

Keeping direct government involvement, and research in government laboratories, out of sight looks like the last concession that will be made to those who oppose SDI. So far, the opposition has in any case been remarkable for its silence. Neither the principal opposition parties nor the media have attempted to make a major political issue out of SDI, despite the extremely emotive nature of defence issues in Japan.

One reason is that the opposition has much bigger problems to think about. In the next few months, Prime Minister Yasuhiro Nakasone will begin the break-up of the state-owned Japan National Railways (JNR) into six regional private companies, plus companies to run the "bullet" train and freight services. The aim is to get JNR's almost unbelievable debt — in excess of the foreign debt of Mexico — off the public shoulders and to shed almost 40 per cent of its third of a million staff. The principal opposition party, the Japan Socialist Party, and the unions have enough to do opposing privatization without worrying about SDI.

Opposition to SDI participation may also arise through factional struggles within the ranks of Nakasone's own Liberal Democratic Party (LDP). The Prime Minister seems clearly to be aiming at calling elections for both upper and lower houses of the Diet in June. At that time, he should be at the peak of his popularity. He will have visited the United States in April and hosted the summit of industrial

nations in Tokyo in June, both of which are certain to heighten his image as Japan's first truly international statesman. If a resounding victory in the elections follows, he will be in a position to call for a change in the LDP rules that will allow him to stay on for a third two-year term as LDP leader, and thus Prime Minister.

Naturally enough, other contenders for the LDP leadership who now begin to see that the end of Nakasone's fixed term of

Yup, we sure could have done with a few of those beauties above Pearl Harbor...



office may be a mirage are less than happy. Among them, Foreign Minister Shintaro Abe is known to be cool towards SDI participation: so internal feuds within LDP could still lead to a change of course.

The only other serious barrier to participation is a 199 all-party Diet resolution banning the military use of space. The resolution has been taken seriously and

has been invoked, for example, to try to stop the self-defence forces using Japanese-launched commercial satellites for military communications. But lawyers already believe that loopholes can be found.

Next week's mission to the United States is plainly aimed at looking for commercial possibilities. It will include engineers from virtually all of Japan's big electronics companies — Fujitsu, Hitachi, NEC, Sony, Oki and Sumitomo — as well as from large self-defence force contractors such as Mitsubishi Heavy Industries. The industrial representatives will split into three groups according to their fields of commercial expertise. They will visit the Lincoln Laboratory of MIT, the Los Alamos National Laboratory and the US Army Strategic Defense Command as well as several industrial companies.

Worries that Japan's lack of anti-espionage laws might limit participation seem also to have been dispelled. US Defense Secretary Caspar Weinberger said last week that Japan would not have to be bound by special confidentiality agreements to carry out SDI research, given the "high level of trust between the two nations".

At the end of the day, however, there is still no doubt that many researchers in industry would prefer to have little to do with SDI. Some, who remember the Second World War, find any involvement with the military regrettable. Others attribute Japan's economic miracle to its refusal to waste talent or money on the development of military rather than consumer technology. They fear they are gradually going to be dragged down into the same hole in which the United States finds itself.

Alun Anderson

Waving the star (wars) and stripes

Washington

No expense was spared, no detail too small at a Washington banquet and award ceremony held last week for heroes of the Strategic Defense Initiative (SDI). Even jaded regulars at defence contractor bashes were flabbergasted when a full-colour guard in revolutionary war uniforms marched to the podium to the beat of fife and drum, bearing the stars and stripes and military flags festooned with campaign streamers.

The Hollywood-style ceremony was held by the American Defense Preparedness Association (ADPA), a defence contractor organization, as part of a two-day classified symposium on SDI at the National Academy of Sciences. The first two recipients of the ADPA Strategic Defense Award were Dr James Fletcher, recently nominated as administrator of the National Aeronautics and Space Administration and chairman of a key early study that supported SDI, and Dr George Keyworth,

a fervent SDI advocate who was until recently director of the White House's Office of Science and Technology Policy. Medallions bearing the ADPA seal were presented to the proud winners by no less than Lt General James Abrahamson, director of the SDI organization.

Banquet guests were then treated to five-minute film clips capturing dramatic high-points of the public-consumption experiment spectacles. The US Army's SDI team and the Lockheed Missile and Space Company took home the ADPA Strategic Defense Technical Achievement Award, for the successful homing overlay experiment in which a dummy missile was intercepted and destroyed in space. Runners-up were Dr Richard Briggs of Lawrence Livermore Laboratory (free electron laser), and Itek Optical Systems Division (atmospheric compensation system, tested on television during a shuttle flight last year).

Tim Beardsley

French elections

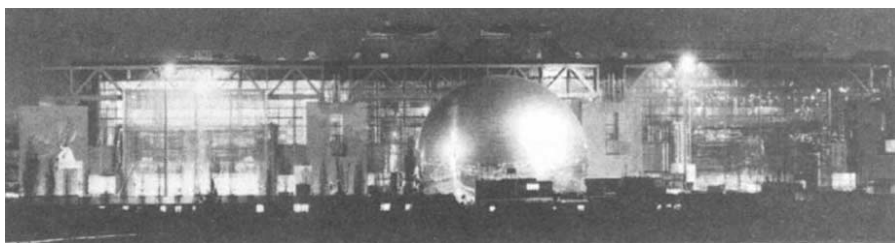
Research loses its ministry

THE French general elections last week, which led to the appointment of a new right-wing French government, have also produced a slap in the face for the research establishment. For M. Jacques Chirac, the new prime minister, has failed to appoint a minister to head the Ministère de la Recherche et de la Technologie (MRT), the nerve centre of the previous government's strategy on research. Last week, staff at the ministry were in limbo, as ministerial appointments were announced which ignored the existence of MRT.

Even worse, the new prime minister, the mayor of Paris and leader of the minority Rassemblement pour la République (RPR) party, has not mentioned research or technology in a list of five priority issues for the future, although the outgoing socialist government would almost certainly have done so. For Chirac, precedence goes to issues such as denationalization, law and order and a return to transferable voting (a system which President Francois Mitterrand threw out in favour of proportional representation in a successful bid to complicate the right's hold on the present parliament).

According to the plans of Chirac's science advisers, responsibility for technology policy and the overall distribution of the French government research budget will be given to a prime-ministerial office one-tenth the size of the present MRT, which would be faster acting and more flexible than the socialists' ministry. But by late last week, no announcement had been made about the setting-up of such a body.

This turn of events presents one big question-mark for Eureka, the programme for European cooperation in high-technology innovation which was created with a great flourish by President Mitterrand last year. Anomalously, Mitterrand remains president until the next presidential elections (which must take place by 1988). Constitutionally, he will retain major powers over foreign and defence policy, and already he has vetoed one of Chirac's nominees for the job of foreign minister — ostensibly because the candidate was too much in favour of the US Strategic Defense Initiative (SDI). Eureka will fall slap into the centre of this turmoil, as the programme was born as a response to the US call for European participation in SDI, and involves foreign, defence and research policy in high degree. Yet the ministry which has been overseeing the programme (MRT) is no longer to exist: conflict between Chirac and Mitterrand on foreign policy seems inevitable.



THE "Cité des Sciences et de l'Industrie", France's giant leap into modern interactive exhibition methods for science and technology, is now open to the public. Originally commissioned by the previous French president, M. Giscard d'Estaing, the Cité has grown out of a gigantic and failed abattoir built to serve the whole of the Paris region.

The redesigned building includes two "domes" on the roof which will follow the

Sun and channel light to the building through a system of mirrors. The 36-metre-diameter "Geode" in front of the building houses a 180-degree projection system and can seat 370 people. The whole complex cost FF4,450 million (around £445 million) and, with 800 staff, will cost FF158 million (\$16 million) a year to run. Parts of the exhibition have yet to be completed, but all should be in operation by September this year. Robert Walgate

As for research, M. Chirac's actions so far are entirely consistent with statements made by his science advisers before the election. The only mention of research in Chirac's new, slimmed-down administration is at the ministry of education, where one of three junior ministers assisting the principal education minister will be responsible for "the universities and research". If the previously advertised programme is carried through, this new "delegate minister", theoretical chemist M. Alain Devaquet, will now take control of the major French research council, the Centre National de la Recherche Scientifique (CNRS), and begin to dismantle parts of it in order to decentralize policy-making and to give university researchers more direct power over their laboratories.

Devaquet himself, however, is less

feared by some of the architects of the previous government's science policy than others Chirac might have chosen. A 43-year-old professor at the University of Paris (South), he studied chemistry at the École Normale Supérieure de St Cloud from 1962 to 1966, and has worked in North America at Cornell University and the University of Western Ontario. He is a member of a respected CNRS "laboratoire associé", one of the many CNRS laboratories jointly supported by the universities and by CNRS. Many such laboratories, however, have a sense of having suffered from CNRS policies of concentration, which are claimed to have favoured the CNRS's fully-owned "laboratoires propres", and this factor may shape some of Devaquet's own thinking.

Robert Walgate

Nuclear energy

Weizsäcker changes his mind

Hamburg

CARL FRIEDRICH VON WEIZSÄCKER, the physicist and philosopher whose brother, Richard von Weizsäcker, is West German president, has changed his mind about the place of nuclear energy in the fuel economy of West Germany. As director of the Max-Planck Institute of Future Research until his retirement in 1980, von Weizsäcker judiciously balanced his opposition to nuclear weapons with the opinion that peaceful nuclear energy is a necessity. But now, in a foreword to a book to be published shortly, he announces that he has changed his mind.

Although now 74, von Weizsäcker's opinion is still influential in West Germany. His revised opinion accompanies a book *Die Grenzen der Atomwirtschaft* (The limitations of the nuclear economy) by Klaus Michael Meyer-Abich, a Hamburg state senator, and Bertram Schefold, a Frankfurt economist.

Von Weizsäcker has been influential in

the West German debate on nuclear energy since at least 1957, when he organized the "declaration of the Göttingen eighteen" in which a number of distinguished West German physicists declared their opposition to nuclear weapons. In 1983, he was the adviser on national security policy to the social democratic candidate for federal chancellor, Hans-Jochen Vogel.

According to the unpublished foreword, von Weizsäcker has changed his mind because of the "sleepless nights" he endured worrying about the problem of the protection of civil nuclear power plants from attack, either by terrorists or in time of war. He says that he has come to prefer that the development of solar energy should now take precedence because of his fear of "outrages", and that he does not dispute the possibility that, in a peaceful world, nuclear energy could again "render mankind an important service".

Jürgen Neffe

Japan's research

Industry imitating life tomorrow

Tokyo

A SMALL team inside Japan's Ministry of International Trade and Industry (MITI) is busy preparing the framework for a major research programme — "the human frontier project" — that will try to create a new interface between basic biological research and technological development. With luck, the programme could be launched during the summit of the leaders of the Western industrial nations in Tokyo in May.

The human frontier programme takes its fanciful name from part of a report submitted to MITI's Agency of Industrial Science and Technology by a special advisory panel. Among its members were Leon Esaki, Nobel prizewinner and head of IBM's Tokyo Research Center. The report, entitled *Technological Development and International Exchange for the 21st Century*, makes the customary criticisms of Japan as a "free rider" on the West's basic research endeavours. It goes on to propose the "human frontier" plan in that curious mixture of officialese and hyperbole, studded with fashionable high-tech words taken from the English, that seems to characterize the workings of small advisory groups. Among other things it calls for "researchers with centripetal force (!) to be collected in one place so that through repeated *burainsutomingu* (brain storming), *tema* (themes) and *conseputo* (concepts) can be clarified . . .". As with other MITI research projects, however, exotic titles and preliminaries are no guide to the seriousness of the contents.

What is being suggested is a long-term (perhaps twenty-year) programme aimed at carrying out basic research in the biological sciences but with very close attention to possible technological applications. MITI research projects, which have always been partly carried out in government research institutes and partly in industrial companies, seem just the format for this approach. And if it seems unlikely that industrial companies would be interested in basic biological research, it is worth remembering that both NEC and Mitsubishi Electric are running invertebrate neurophysiology programmes on the grounds that better computers can be built by studying the way living organisms transmit and process information.

That idea is behind one of the three main areas proposed for the human frontier programme: an attempt to imitate functioning of brain and nerve networks to find new ways of information management and control circuitry that could be implemented in new kinds of electronic materials. A second area is the imitation of chemical processes taking place within

living things. The aim is to build better industrial catalysts, carry out reactions using less energy and to trap light energy with artificial photosynthesis systems. Imitation of muscles and sinews of living organisms comes third, with the aim of building energy-efficient "physiological machines". The overall philosophy is of basic research for the sake of its applications, contrasting with the popular Western concept of basic research for the sake of "the advancement of knowledge".

The human frontier programme is not the only new project to be aired in recent weeks. A separate private advisory panel to the Prime Minister, the Advisory Group on Structural Economic Adjustment for International Harmony, is expected shortly to call for a Japanese

version of the European Eureka high-technology development project. Artificial intelligence, new materials, optoelectronics and robotics are likely to be its key themes and international collaboration is to be invited.

What in part makes thoughts of such projects possible is the dramatic drop in Japan's oil bill. Two-thirds of Japan's energy needs are met by oil and every drop of it is imported. The halving of crude oil prices (not yet passed on to the consumer) gives the government an unparalleled opportunity to spend money in new ways. A second factor is the problem of Japan's participation in the US Strategic Defense Initiative (SDI) (see page 294). Public opposition to military research in Japan is strong and it is thought that major, non-military international research might sweeten the involvement in SDI that the government is clearly aiming at.

Alun Anderson

Japanese space programme

Abundance of satellites planned

Tokyo

NEW and revised plans for Japan's space programme were announced last week by the Space Activities Commission (SAC). Altogether the launch of some eleven satellites plus a couple of space experiments have been fixed for the next five fiscal years, along with development of new and more powerful launch vehicles.

SAC is a part of the Prime Minister's Office responsible for overall coordination of the space programme's two arms: the scientific programme run by the Institute of Space and Astronautical Science (ISAS) and the applications programme run by the National Space Development Agency (NASDA). On the academic side, the chief news is that a magnetosphere observation satellite "Geotail" will be launched by the end of the 1990 fiscal year. The satellite is due to be launched by the space shuttle in a cooperative programme with the United States and will make observations on the structure and dynamics of the long magnetic tail formed on the night side of the Earth.

In the nearer future, two other Japanese experiments are due to be taken up by the space shuttle. No changes in their dates have yet been announced by the US National Aeronautics and Space Administration as a result of the explosion of the Challenger. From ISAS come Space Experiments with Particle Accelerator (SEPAC). The idea is to inject charged particles from electron and ion beam accelerators into the Earth's ionosphere and magnetosphere to generate an artificial aurora which will aid understanding of the behaviour of space plasma and the generation of auroral lights. NASDA aims at a more practical application with

its First Material Processing Test, to be carried on Spacelab. Various experiments for producing materials under gravity-free conditions are under development.

Three other important missions will come from ISAS. As had been expected, Astro-C will be launched this fiscal year to continue ISAS's highly successful series of X-ray observation satellites. Heavier and bigger than its two predecessors launched in 1981 and 1982, it should be able to observe X-ray sources in more distant galaxies as well as in the Milky Way. Following that, Exos-D will be launched in fiscal year 1988 to continue the series of satellites observing the upper atmosphere and aurora. Next, in fiscal year 1989, the "Muses" interplanetary probe programme will begin with the launch of Muses-A, a satellite designed to test the orientation, control and data-transmission systems needed to make possible a flight to the Moon.

The SAC plan also reveals that NASDA's H-II rocket should be ready for its maiden flight in fiscal year 1991. Using almost entirely home-developed technology, it will have two liquid hydrogen/liquid oxygen stages and be capable of putting a two-ton satellite into geostationary orbit. But before that its predecessor, the H-I, capable of handling only a 550-kg payload, should, between fiscal years 1987 and 1989, have launched six satellites. Success will of course depend upon the conclusion of the current development of the H-I's second-stage liquid hydrogen/liquid oxygen engine. Before H-I comes into operation, one more satellite will be launched by the last of the N-II liquid fuel rockets, the Mos-I marine observation satellite.

Alun Anderson

US research costs

Pressure builds in Congress

Washington

RESPONDING to howls of indignation from US universities and Congress, the White House's Office of Management and Budget (OMB) now says it is willing to delay implementing new regulations for compensating universities for the indirect costs associated with federal grants.

Joseph Wright, deputy director of OMB, made the offer to postpone implementation until 1 July at a meeting of the House of Representatives subcommittee on science research and technology to consider the planned changes. Members of the subcommittee were dismayed by OMB's "haste" to make sweeping changes to Circular A-21, the document providing the blueprint for cost reimbursement. The proposed revisions would limit the reimbursement for administrative costs to a fixed percentage of the total amount of a grant; 26 per cent in 1986, dropping to 20 per cent in 1987.

OMB estimates the move will save \$300 million over the next two fiscal years. OMB published its proposed changes in the *Federal Register* on 12 February, with a comment period of only 30 days. The revisions were to take effect from 1 April, although individual agencies were allowed to delay implementation for up to a year, and most had said they would need at least until 1 July to change their procedures.

Even before the new regulations were announced, the university community was seething over rumours of OMB's plans (see *Nature* 319, 346; 1986). The complaints were that OMB was making the changes without consulting those affected, and that the 26 per cent figure was arbitrary, not reflecting varying needs in different institutions. When the changes were finally announced, universities were indignant at the short period allowed for comment, complaining that OMB could not be very interested in what they had to say.

Since 12 February, university presidents and their allies in Congress have unleashed a flood of letters to OMB protesting the changes. Forced into a conciliatory posture before the subcommittee, Wright was nonetheless unwilling to extend the 30-day comment period. He did, however, agree to meet Dale Corson, chairman of the Government/University/Industry Roundtable, to establish an agenda for discussions in the coming weeks. In addition, subcommittee chairman Doug Walgren (Democrat, Pennsylvania) indicated that his committee would continue to take testimony on the amendments, and would require OMB to respond to them.

Joseph Palca

British astronomy

Time for Greenwich to move

A DECISION has finally been made about the fate of the two ground-based observatories in Britain. The Science and Engineering Research Council (SERC), which supports the Royal Greenwich Observatory (based, since 1948, at Herstmonceux in Sussex) and the Royal Observatory, Edinburgh, announced last week that the Royal Greenwich Observatory (RGO) will be moved yet again. A final decision will be made in June between the three sites now under consideration: "in order of priority", Edinburgh and the Universities of Cambridge and Manchester.

The need for some decision arises from the steady replacement of British-based optical instruments by more advanced equipment at overseas observatories, chiefly at La Palma in the mid-Atlantic and Hawaii. Although the two British observatories have provided technical and managerial support for the overseas observatories, the role of RGO in this connection will be much reduced from 1990, with the completion of the 4.3-m John Herschel reflector at the La Palma observatory. Although SERC plans to decide on the eventual location of RGO at its meeting in June, the move will not take place until 1990.

Plainly, SERC has found its decision about the location of RGO exceedingly difficult to make. The need for some reorganization has been apparent for some five years, and has been an urgent issue since November 1984, when SERC put in hand a reappraisal of its pattern of spending. But a working party under the previous chairman of SERC, Sir John Kingman, which produced a list of options for the council in January of this year, was apparently unable to decide between them. It will not be possible to confirm rumours that the Kingman group recommended a continuation of the status quo because, according to Professor E.J.W. Mitchell, now the chairman of SERC, its report will not be published.

At the same time, the possibility of British withdrawal from the Anglo-Australian Telescope has been postponed, at least until 1990. This became a live issue roughly a year ago, when continued British collaboration at the Australian site, which costs SERC £1.5 million a year, was given a low priority by the Astronomy and Space Board of SERC. Australians were quick to point out that the collaboration is regulated by a formal treaty between Britain and Australia with no break clause. Now, it seems, discussions are under way with other potential partners, including Japan, which may lead to a change of status for the Anglo-Australian Telescope, but not before the end of the decade.

SERC's intention is that the move of

RGO from the Herstmonceux site will be self-financing, at least so far as capital costs are concerned. The chief, but still uncertain, element in any such calculation is the value of the fifteenth century castle which dominates the observatory's site in Sussex. One imponderable is the likely value of the building on the open market. Another is whether SERC would be allowed by the British Treasury to keep the whole proceeds of a successful sale.

According to Mitchell last week, the intention is that RGO should retain its separate identity even when it has moved to another site. To the extent that both RGO and Edinburgh provide instrumental and engineering support for overseas telescopes, the two large optical instruments at La Palma and the infrared and millimetre-wave telescopes in Hawaii (the second of which is now nearing completion), the advantages of a merger at Edinburgh are plain. But SERC is also aware of the benefits of siting RGO at a university where astronomy is already strong. One of SERC's disappointments in the present arrangements is plainly that the connection between RGO and the University of Sussex, 20 miles away, is "not the kind of interaction we are looking for", according to Mitchell.

For the time being, there are no firm plans to save recurrent costs by the proposed move, although the Astronomy and Space Board is already committed to reducing its expenditure by £1.5 million a year in the interests of SERC's flexibility. Steps have already been taken to consult trades unions representing staff at RGO.

Among the complications of the move now foreseen are the need to continue the work of the Nautical Almanac Office to which RGO now makes only a nominal contribution. The work of the laser-ranging unit based at RGO will be continued if suitable arrangements can be reached with the Ministry of Defence and with the Department of Trade and Industry. SERC seems conscious of the need to preserve and make accessible the archive of RGO, but not necessarily at the site to which RGO as a whole will move.

By the time the eventual location of RGO is decided in June, more may be known of the long-term relationship between SERC and the National Space Centre established last year by the British government, but still largely a paper entity. The possibility that SERC might become one of the contributors to this organization, which would be used as a vehicle for carrying through space research projects now under the wing of the Astronomy and Space Research Board, is being considered, but negotiations have not yet been completed. □

US Office of Naval Research

More money means more trouble

Washington

THE US Office of Naval Research (ONR) celebrates its fortieth anniversary this year, but its biological research managers have their eyes to the future. The \$40 million spent this year by ONR's life sciences directorate supports some of the most long-term basic research in the federal budget. But as the federal budget for military research has increased in recent years in relation to the civilian budget, the role of ONR seems to be changing. It plans to support more training and education as well as research. Some academics have misgivings.

Almost all the life sciences research supported by ONR (about 15 per cent of the total) is carried out under contracts with universities. ONR uses the newest techniques of molecular biology and biotechnology to produce protein adhesives, deuterated lubricants and other surfactant molecules. Priority research topics include enzymes from organisms that inhabit extreme environments, the interplay of the central nervous system with immune function and the modelling of neural circuits with the aim of understanding "biological intelligence".

ONR selects among research proposals at least partly by independent review. All the bioscience research is unclassified. But there the similarity with a civilian research agency ends. In the words of Dr Steven Zornetzer, associate head of ONR's life sciences directorate, research supported by ONR "should have, or could have down-the-line" applications for the US Navy. That condition is one of the reasons behind a recent controversial ONR decision to close its Naval Biosciences Laboratory at Oakland, California. It has also caused dissent at the Massachusetts Institute of Technology (MIT), where biology faculty have voted once to refuse up to \$4 million per year of ONR support for a new biotechnology training programme.

The Naval Biosciences Laboratory, managed by the School of Public Health of the University of California (UC) at Berkeley, had a total budget of \$5.7 million in 1985 (mostly from ONR) and a staff of about 100. Research there focuses on gene expression in different systems and on "slow" viruses. In 1984, the School of Public Health and the laboratory together inaugurated a molecular parasitology research group, headed by acting laboratory director Dr Nina Agabian, which grew rapidly and now represents about half of the laboratory's staff. But, despite "extraordinarily high marks" at a May 1985 external review, Agabian says that the laboratory was told abruptly last September that the site must be vacated by September 1987. Feelings are still running high.

ONR cites two reasons for the decision. First, the converted Second World War barracks that house the laboratory are outdated, and modification would not be cost effective. And, second, molecular parasitology with potential medical applications should be the responsibility of the US Army rather than ONR in accordance with an inter-service agreement that the army should support infectious disease research. Agabian's group now expects to be rehoused in a new building as part of an inter-campus parasitology effort with UC San Francisco. The relationship with ONR seems over.

The navy clearly had reason other than mere lack of interest in the laboratory's dominant research area for its decision. The laboratory's contractual position is unusual (although not unique). Other non-competitive research contracts were modified at the same time.

Worries that basic research may be skewed towards military objectives underlie in part the vote by MIT's biology faculty last month not to apply to ONR for support for its planned biotechnology training programme. ONR announced last January that (together with the Defense Advanced Research Projects Agency) it proposed to spend an additional \$25 million in fiscal year 1986 on 5-year block research grants worth up to \$4 million. The increase represents ONR's share of the Pentagon's University Research Initiative, which comes on stream this year. One of the eligible grants, for marine bioengineering, would support graduate student research training in molecular biology of marine organisms, with special reference to bio-fouling and drag reduction.

MIT's plans for an interdisciplinary biotechnology training programme had been held up by lack of funds so that the new ONR grant seemed a godsend. But the biology faculty objected to applying (albeit at a poorly-attended meeting), although the other faculties that would be involved have raised no objection. The biology faculty is expected to reverse its decision in a second vote this week.

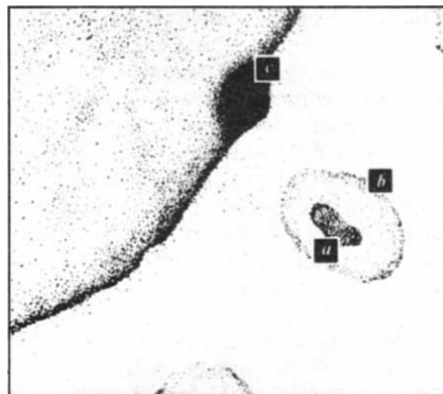
ONR already supports some researchers in the MIT biology faculty, but opponents of ONR support distinguish between individuals who accept specific ONR research grants and MIT as a whole accepting that a whole course should be supported ultimately by the Department of Defense.

Dr Maury Fox, chairman of the biology department, who does not oppose the idea that ONR might support the course, would also prefer educational programmes funded by non-military sources.

No help on AIDS

THE illustration below is the only one in a full-page advertisement headed **ARE YOU AT RISK FROM AIDS?** which was placed by the Department of Health and Social Security in British newspapers last week. Its caption runs: "a, AIDS nucleoid containing the biological message to cause damage (*sic*). b, Lipid membrane (very fragile). Packages virus and allows movement between cells. c, T helper cell/white cell."

The text of the advertisement, which is admirably expressed in plain English, does not mention nucleoids, lipids, virus packaging, T helper cells or white cells. Asked



why such an enigmatic illustration and confusing legend should have been allowed to mar an otherwise excellent message, a spokesman for the department said it was included to relieve the monotony of the text. Asked what a nucleoid was, she came back with the definition "a granular or fibrillar substance in certain erythrocytes which resembles a nucleus". That is indeed one dictionary definition (for example *McGraw-Hill Nursing Dictionary*). But the appropriate definition is "A term used by electron microscopists to describe the electron-dense centrally placed region observed in certain viruses" (*A Dictionary of Virology*, Blackwell). □

Fox believes, however, that that is a matter for national policy and "I'll have to live with it". Others are unconvinced. Frank Solomon, assistant professor of biology at MIT, opposes ONR block grants and believes the Pentagon would not hesitate to classify "biosludge" research that turned out to have important weapons applications. And he is opposed in principle to the military supporting and controlling a larger proportion of US basic research, selecting topics on other than exclusively scientific merit.

Others fear that the "increased collaboration with ONR" mentioned with the grant prospectus might increase opportunities for military direction of the work. Whatever the outcome of the MIT debate, similar questions may be asked on other campuses as defence research dollars spread further.

Tim Beardsley

Medical Research Council

Government blamed for decay

FORCED to lay bare its plans for the next five years, the British Medical Research Council (MRC) has identified several areas of research into which it intends to pump money but is much less categorical about where the money will come from.

MRC's first "Corporate Plan", published last week, has been produced in response to a request by ministers and the Advisory Board for the Research Councils for each council to explain how it intends to cope with the decline in the overall funds for research that are forecast for the next five years. In a grudging introduction to its plan, MRC asks how it can plan for five years when its budget is the result of an annual bid with an uncertain outcome, made worse by late decisions.

In general terms, the plan states that MRC cannot simply opt out of some areas of medical research. So it will have to raise the threshold for providing funds in all areas of research. At the same time, there is a need for concentration of resources within certain areas in order for the research to be internationally competitive.

The one such area of research that is clearly identified in the plan is neurobiology. This comes as no surprise, since MRC has repeatedly stressed that British neurobiology is in a poor state but has failed to do much about it. "There is a need for a large centre where the development, interactions and molecular biology of nerve cells and their supporting com-

ponents can be studied... in order to retain in this country a number of our most gifted workers", says the report. Additionally it identifies an urgent need to establish at least one major new centre in Alzheimer's disease and a large new facility of parasitology, particularly since spending on tropical medicine in Britain "can only be described as derisory".

Other priorities include a new MRC unit to study disorders of movement and balance, an expansion of support for protein engineering and a need for strong centres for the clinically led study of infectious diseases.

The problem is how to support the new initiatives. MRC uses its plan to say that while remaining alert to ways of increasing its income from commercial resources — which accounted for less than £2 million of the £123.5 million income last year — the generation of income will not become a guiding purpose of its research programmes. Nor does the council intend to become a fund-raising organization.

Even without the need to finance new initiatives, the plan identifies the need for extra spending. First, the salaries that can be offered are already uncompetitive in some key areas of research. Second, there is a pressing need to increase the number of research studentships and training fellowships to avoid the loss to research of many gifted graduates. And, third, there is the severe problem of outdated equip-

ment in MRC research establishments.

To meet all these demands, MRC is to impose higher standards in judging whether to continue an existing research unit upon the departure of its director, and to encourage all directors not to replace staff who leave. It intends also to introduce compulsory retirement "on grounds of reduced efficiency" for its scientists and to be tougher in assessing the cost-effectiveness of research it supports.

Lacking any real solution to its financial problems, however, MRC uses its corporate plan repeatedly to attack government policy. Reduction of public expenditure is a legitimate aim, it implies, but "the cuts have been made too rapidly for any sensible accommodation... there has been little appreciation of the damage this is causing to the nation's research capability".

Peter Newmark

West German environment

Greens bite back in Hesse

Hamburg

HERR Joschka Fischer has begun to make his mark as the first member of the environmentalist Green party to hold office as a regional government minister. At the beginning of the month, Fischer made public his first set of regulations, intended to improve water quality of the rivers Main and Rhine, which will bear most directly on the operations of Farbwerke Hoechst, the largest chemical manufacturer in West Germany.

Fischer owes his post as a minister in the Hesse government to the failure of the Social Democrats (SPD) to retain their majority in the Hesse parliament. The Hesse Greens were at first divided over the proposal that they should form a coalition with the SPD, but eventually a "red-green" alliance was formed.

Far from being a stranger to politics, Fischer won acclaim during his two years in the federal parliament (Bundestag) as an eloquent speaker. He was required to retire from the Bundestag when the Greens decided that their representatives would be rotated every two years.

In fact the pollution limits so far decreed are likely to cause Hoechst very little trouble. Apparently relieved that the regulations are not more worrying, the company says that it will need no new equipment to keep its daily discharge of mercury below 1.125 kg and cadmium below 0.6 kg (a fourfold reduction in each case). The daily discharge of all acid is to be reduced from 150 to 10 tonnes. However, Fischer does not conceal his intention, in collaboration with Hoechst, that pollution should be further reduced when the new regulations expire in 1987.

Jürgen Neffe

British farm research

Downbeat plan for agriculture

THE British Agricultural and Food Research Council (AFRC), which has seen the British government wash its hands of responsibility for agricultural research in the past five years, has now produced a corporate plan for the years ahead which is reciprocally dismissive of what it can expect from government. The chief message is that, within a budget that will have shrunk by some 26 per cent (if inflation works out at an average of 5 per cent a year) between 1983 and 1991, the council will somehow manage to increase its spending on university research grants to more than £8 million a year.

The council goes further than its predecessors to acknowledge why the British government has been beastly towards AFRC by noting that the success of European agricultural research, which is manifested by the large surpluses of commodities accumulated under the Common Agricultural Policy, are a proof that merely increasing production is no longer a sufficient objective for research. But the plan says that while consideration such as

the nutritional value of the public diet are now of greater importance than in the past, it remains the case that the production of high-value crops at low cost is an important objective of research.

The plan also marks out, as a new direction for its research, the better integration of agriculture with environmental considerations, an issue raised two years ago by a report of the House of Lords Select Committee on Science and Technology which was generally critical of farmers.

The corporate plan shows that most areas of AFRC research will decline in the years ahead, but that there will be a growth (to 14.2 per cent by value) of food research. It says that it hopes to raise funds from other than government sources to increase what it has available for research, but is plainly uncertain how to calculate what may be available given the government's intention to support a substantial part of the cost of British agricultural research by levies on those sections of the farming industry that can be made to pay for the service. □

Polish agriculture

Plans for foundation impeded

POLAND'S controversial "agricultural foundation", sponsored by the Roman Catholic Church, which plans to raise money in the West to inject new technology into Polish agriculture, has run into further trouble. The original scheme would have benefited only private farmers, since the 25 per cent of land in the "socialized" sector has for many years received massive government investment. But Polish sources now report that the government is pressing for a proportion of the fund to go to socialized agriculture.

Since the idea of the foundation was first mooted four years ago, government negotiators have shown every sign of trying to block progress. Meetings with Church representatives have been postponed on every conceivable excuse, ranging from the "wars of the Crosses" (protests by secondary school children about the removal of crucifixes from their classrooms) to the murder of Father Jerzy Popieluszko.

Even after the enabling legislation for the foundation was passed in autumn 1984, the wrangling continued. Under the terms of the foundation, an initial \$28 mil-

lion was raised for pilot studies, and ten projects were worked out, including water development, agricultural field services, milk production and tractor repairs. But the government was reluctant to let work begin on any of these until the whole cost of the planned rescue operation (\$1,800 million over five years) had been raised.



lion was raised for pilot studies, and ten projects were worked out, including water development, agricultural field services, milk production and tractor repairs. But the government was reluctant to let work begin on any of these until the whole cost of the planned rescue operation (\$1,800 million over five years) had been raised.

Another cause of dispute was the foundation's plan to import agricultural machinery and supplies, purchased for hard currency, and to sell them to farmers at current market prices, but for zloty, partly on a deferred payment basis. (The zloty

fund so created would eventually be used for in-country investments such as artesian wells and for improving the social infrastructure in rural areas.) The Church wished the imports to be duty-free; the State pressed for the payment of the full rate of duty. Shortly before the parliamentary elections last autumn, the government seemed prepared to concede this point, with the unspoken *quid pro quo* that the Church would not encourage the calls for a boycott of the elections. When the bishops themselves stayed away from the elections, the attitude of the government hardened perceptibly.

By the beginning of this year, a new point of contention was raised. The gov-

ernment hinted that the \$200,000 pledged to it by Mr Lech Walesa (his 1982 Nobel peace prize winnings) would "politicize" the fund irretrievably. A face-saving formula was worked out by which the organizing committee "relieved Mr Walesa of the moral obligations he undertook when he donated the prize to the foundation" and Walesa himself decided to withdraw his winnings.

A Vatican Radio communique which announced the organizing committee's decision at the end of February was couched in terms suggesting that, once this matter is resolved, the foundation could implement the pilot projects. But the new demand that "socialized agriculture" should have a share of the fund suggests that the Polish government is still set on prolonging the negotiations as long as possible.

Vera Rich

US agriculture

More food, but fewer farms

Washington

THE new biotechnology should allow the United States to continue meeting domestic food requirements and to contribute significantly to world food demand over the next 20 years, according to a new report* by the Office of Technology Assessment (OTA). But the same technologies could also exacerbate a trend towards fewer and larger farms, unless Congress takes action to prevent it.

OTA estimates that a 1.8 per cent annual increase in world food production is needed to meet demand by the year 2000, and 83 per cent of this increase has to be from increased production through new technology. OTA agrees with other studies that animal production is likely to be the first to benefit from genetic engineering and other new techniques; plant biotechnology will also benefit, while information technology should allow improved crop and herd management.

Under the most likely assumptions, including application of new technology, OTA estimates that food production could increase by 2 per cent per year. But if the utilization of new technology were significantly less than that OTA thinks most likely, the goal necessary to meet world demand would not be met.

The very detailed report appears to have received a favourable reaction in the US Department of Agriculture. But its most controversial conclusions — because they have immediate consequences for domestic policy — are those relating to different types of farms. OTA argues that very large farms (annual sales of more than \$250,000) will be well able to apply the new technologies by themselves, whereas moderate and small farms will be less able to do so without targeted help from the government. As a result, the

relative competitive advantage enjoyed by large farms will increase as technology advances. OTA predicts that the number of farms will fall from 2.2 million in 1982 to 1.2 million by the year 2000. The number of small farms (less than \$100,000 annual sales) would fall from 60 to 50 per cent.

Some have questioned this prediction: Dr William Brown, chairman of the National Academy of Sciences' Board on Agriculture, believes that the rate of disappearance of small farms is exaggerated in the report.

The report is timely: Congress is in the process of considering the administration's proposed agriculture research budget for the coming fiscal year. Research and development in the Department of Agriculture would (in dollar terms) be hit harder under the proposed budget than almost any other department; however, the reductions are concentrated in the extension services and in programmes targeted to local problems, which the administration has long held should be paid for by local and State governments. Competitive grants in the agricultural research service — which addresses long-term research — are protected. Congress has previously resisted attempts to cut local agricultural applied research, and the House of Representatives' subcommittee dealing with appropriations for agriculture last week made clear its concern about the extent of proposed reductions: some programmes would take cuts of 60 per cent, others would disappear completely. But with the Gramm-Rudman deficit reduction act in place, all bets are off over how Congress will act.

Tim Beardsley

*Technology, public policy and the changing structure of American agriculture. Office of Technology Assessment, 1986.

Soviet education

Helping towards better careers

THE Soviet Union has introduced new rules for admission to higher education, designed to make university training more job-orientated. Instead of students being assessed simply on the sum of marks gained in their entrance examinations, an additional set of marks will also be awarded on the basis of interviews. The interview panel will assess each applicant's chances of success in his or her proposed field of study, and will provide guidance as to career choice. Commenting on the new procedures, the Moscow daily *Pravda* noted that they will give a clear advantage to young people who have begun to prepare themselves for their future profession while still attending secondary school.

The new admission rules are thus a continuation of the policy, introduced in the general education sector two years ago, of making study more closely related to future work.

In practical terms, the new rules will reduce the work-load on the entrants. Instead of having to offer four subjects, they will now only have to offer three: Russian language and literature (non-Russians may offer their native language if that is to be their language of tuition), a choice of mathematics, biology or social science (according to the student's chosen field) and the subject that the student eventually hopes to pursue.

Furthermore, it will be easier for would-be entrants to make applications to more than one university or institute. Until now, except for a handful of elite Moscow institutions, which held their entrance examinations in July, all examinations took place during August. As the Soviet Union has no national clearing-house system, this meant, effectively, that students keen to secure a place would aim at an institution rumoured to be easy on admissions. (This was particularly true of young men wanting to defer military service until after graduation). As a result many of the more cautious applicants set their sights rather lower than was necessary.

Now, according to *Pravda*, all higher education establishments in Moscow city, Moscow *oblast'* and Leningrad will, "as an experiment", hold their entrance examinations in July. This move, *Pravda* pointed out sententiously, "will allow those who did not pass the entrance examinations to one of the Central higher educational institutions to apply their efforts to gain entrance to a higher educational institution in some other town".

Vera Rich

Environmental pollution

US suit charges cancer causation

Washington

In what may become a landmark personal injury case, the families of seven childhood cancer victims from Woburn, Massachusetts, are suing two large US corporations said to be responsible for the illnesses by polluting the local drinking water. Some 40 expert witnesses have been lined up by both sides to argue whether an excess of acute lymphocytic leukaemia and other conditions in east Woburn could be due to halogenated hydrocarbons that were found in two polluted wells in 1979.

The case is one of the first in which it has been claimed that specific chemicals in drinking water caused specific illnesses. The wells, whose water mainly supplied east Woburn, tapped a different aquifer from Woburn's other wells. They were first used in 1967 and were closed in 1979 when trichloroethylene and tetrachloroethylene were found in the water at 200 parts per billion (10^9) and 20 parts per billion respectively; other pollutants included (1,1,1)-trichloroethane, (1,2)-*trans* dichloroethylene and chloroform. The plaintiffs claim that the chemicals came from dumping or accidental spills at nearby plants owned by W.P. Grace and Company and Beatrice Foods. A third company, Unifirst, a dry-cleaning company, has already settled for \$1 million.

A local clergyman apparently first drew attention to a suspected leukaemia cluster in east Woburn, and the town's high incidence compared to national rates was subsequently confirmed in a study by the Massachusetts department of health. Nineteen cases were found where only six would have been expected; and renal cancer was also significantly elevated. The department concluded, however, that the "hypothesis suggesting that the increase in leukaemia incidence was associated with environmental hazards in Woburn . . . is neither supported nor refuted by the study findings".

But a more recent study by S.W. Lagakos, B.J. Wessen and M. Zelen of Harvard School of Public Health and Dana-Farber Cancer Institute does find a positive statistical association between access to water from the polluted wells and childhood leukaemia, as well as perinatal deaths and some congenital abnormalities. This study, in press with the *Journal of the American Statistical Association*, does not compare incidences in Woburn with those elsewhere, but establishes that within Woburn the diseases are found predominantly among those who drank more of the polluted water.

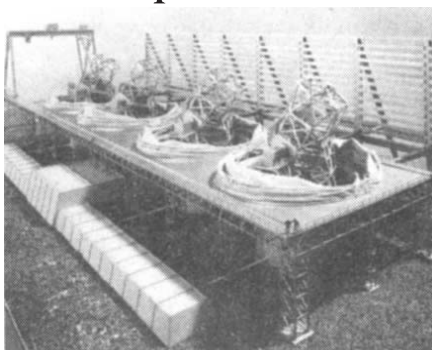
One of the pollutants, trichloroethane, is said recently to have been shown to cause leukaemia in rats. But a major

weakness in the plaintiffs' case is that nothing is known about the nature and degree of pollution of the wells at any time other than 1979, when they were immediately closed. The known carcinogens that were then in the water supply were present in far lower concentrations than have been shown to cause cancers in laboratory animals and in human occupational health studies. But it is argued that the fetus may be more susceptible to carcinogens than the adult, and that plaintiffs are basing much of their case on recent immunological tests indicating abnormalities in the blood of the plaintiffs' families.

The trial, which is in its third week, is expected to last for several months. First it will be decided whether the defendants contributed to groundwater contamination; medical questions will be examined only if the answer to the first question is yes. Beatrice Foods is refusing to comment on detailed aspects of the litigation, although it is pointed out that the accusation is that the company allowed pollution of ground it owns adjacent to its plant. W.P. Grace says that its plant, which manufactures stainless steel food processing equipment, may have leaked small quantities of cleaning agents and degreasers onto its own property, but that investigations (including drilling 50 test wells) had shown this would not have found its way into the wells.

Tim Beardsley

Europe's optical telescope for 1990



STILL just a model, this is the 16-metre optical telescope, the "Very Large Telescope", planned for the European Southern Observatory (ESO) in Chile. ESO technical committees are now meeting to establish exact specifications for the new instrument with the aim of presenting plans for a DM 300 million project to ESO member countries in time for work to start in 1987. The VLT is to be sited at an altitude of 2,400m, and the human figures in the above photograph put its size into perspective.

Sakharov's achievements

SIR—Ernest B. Gliner's letter "Another reason for saving Sakharov" (*Nature* 318, 513; 1985) notes only his work on gravitation. His major contributions in particle physics, fusion research and cosmology provide additional evidence that Sakharov's scientific contributions will be remembered, like those of Galileo, long after the names of those who persecuted him are forgotten. Perhaps the new Soviet leaders will note that their grandchildren may study the work of this great Russian physicist and wonder about the role of their grandfathers in his persecution.

In 1950 Sakharov proposed a solution to fusion's major problem: how to contain the reaction at its enormously high temperatures where no known materials can survive. Sakharov's "magnetic bottle", called the "tokomak" (ref. 1, p.49), which uses magnetic forces to contain the hot fusion fire, is currently the front-line approach to fusion reactors and may go down in history as man's answer to the energy crisis.

Sakharov's two remarkable 1966 papers on particle physics were far ahead of their time. Sakharov and Ya. B. Zeldovich (ref. 1, p.271; ref. 2) obtained relations between masses of different known particles in surprising agreement with experiment by assuming the new quark theory of matter which built all these particles from the same three basic building blocks arranged in different combinations, and held them together by forces which, although unknown, were always the same in different particles. Sakharov's work was ignored because the physics establishment did not take quarks seriously at that time.

Even more remarkable was Sakharov's resolution in 1966 of the apparent contradiction between the "big bang" theory of the origin of the Universe and the failure of astrophysical observations to reveal any trace of antimatter in the Universe. In his theory, the antimatter created in equal amounts with matter in the big bang decayed more rapidly than matter, leaving the observed excess matter. Sakharov demonstrated (ref. 1, pp.147, 151) three necessary conditions for this mechanism: (1) the recently discovered CP violation asymmetry between the interactions of particles and antiparticles; (2) a departure from thermal equilibrium; (3) violation of the law of baryon number conservation. This required the proton to be unstable and to decay into electrons and mesons, but so slowly that proton decay would not have been detected in any experiments.

His theory was not accepted because of a number of assumptions ridiculed as crazy at the time: (1) the existence of new very heavy boson particles; (2) the existence of quarks and of new interactions between quarks, electrons and these new

heavy bosons; (3) violation of the sacred principle of baryon conservation.

Today Sakharov's theory is the accepted view on the antimatter problem, his crazy assumptions are now a central part of the standard theories, and his three conditions are accepted as necessary for any cosmological model. Quarks, new heavy bosons which allow protons to change into electrons, violation of the law of baryon conservation and predictions that the proton must decay are all essential ingredients of the new "grand unification" theories. Many large expensive experiments are searching for Sakharov's predicted proton decay.

Even in his isolation in Gorki he has managed to continue the development of his 1966 ideas into new work on the quark theory of matter³ and the cosmology of the early Universe⁴. His continual hopes for freedom in the Soviet Union and an end to the nuclear arms race seem like wild dreams. But we can all hope that once again Andrei Sakharov is right and only 10 years ahead of his time, and that the new Soviet leaders will allow him to return to Moscow with free access to scientific libraries, institutes and personal contacts so that he can pursue his work freely for the benefit of all mankind.

HARRY J. LIPKIN

High Energy Physics Division,
Argonne National Laboratory,
9700 South Cass Avenue,
Argonne, Illinois 60439, USA

1. Sakharov, A.D. *Collected Scientific Works* (eds Ter Haar, D., Chudnovsky, D.V. & Chudmovsky, G.V.) (Dekker, New York, 1982).
2. Lipkin, H.J. *Ann. N.Y. Acad. Sci.* **452**, 79 (1985).
3. Sakharov A.D. *Zh. eksp. teor. Fiz.* **78**, 2112 (1980); **79**, 350 (1980); *Soviet Phys. JETP* **51**, 1059 (1980); **52**, 175 (1980).
4. Sakharov, A.D. *Zh. eksp. teor. Fiz.* **79**, 689 (1980); *Soviet Phys. JETP* **52**, 349 (1980).

Open house

SIR—In your article "What price freedom?" (*Nature* 319, 608; 1986), you say that "IUPPS is affiliated to a UNESCO umbrella organization which advocates an academic boycott of South Africa". This point needs to be clarified. The International Union of Prehistoric and Protohistoric Sciences (IUPPS) is affiliated to the International Council for Philosophy and Humanistic Studies (CIPSH). CIPSH has considered itself bound by UNESCO policy and will not give funds either for activities taking place in South Africa or to persons falling under the jurisdiction of South Africa.

CIPSH, however, believes that international congresses should be open to all bona fide scholars, irrespective of domicile, race, religion or belief. This policy was set out in an exchange of correspondence between the secretary-general of CIPSH and the secretary-general of

IUPPS in August 1985, specifically in response to Professor Ucko's enquiry to the latter as to the official attitude of IUPPS and CIPSH towards the participation of persons working in South Africa in the activities of IUPPS. Moreover, the CIPSH bureau, at its meeting in Istanbul about the beginning of December 1985, reiterated this CIPSH policy and indicated that the organizers of the Southampton congress should take every precaution to ensure the free participation of all in their activities.

Failing such arrangements, the bureau resolved, the organizers of the congress would be entitled to cancel the holding of the meeting, since it would not be able to ensure completely free participation and academic freedom.

In this attitude, then, it does not seem that the policy of CIPSH "advocates an academic boycott". On the contrary, its policy towards international congresses seems to square with that of the International Council of Scientific Unions (ICSU): both parties favour the free circulation of scientists and an "open house" participation in international congresses.

PHILLIP V. TOBIAS

Department of Anatomy,
University of the Witwatersrand
Medical School,
York Road, Parktown,
Johannesburg 2193, South Africa

Too many fourth states of matter?

SIR—In the short period of time between May and October last year, I see that the 'fourth state of matter' was referred to in *Nature* in two book reviews, one on plasma physics¹ and the other on polymers². The same phrase has been used in connection with the discovery of icosahedral symmetry.

Physicists should realize that if there is such a thing as the fourth state of matter, there can be only one of it. Moreover, physicists should be aware that in 1940 the attribution was used to describe the superfluidity of helium-3 (refs 5, 6). But the phrase has also been used to describe other kinds of superfluidity such as in the superconducting state of electrons⁷.

Is there a danger that this practice will get out of hand?

J.M. GOLDSCHVARTZ

Clavecimbellaan 273,
2287 VK Rijswijk zh,
The Netherlands

1. Lashmore-Davis, C.N. *Nature* **315**, 162 (1985).
2. Eisenberg, H. *Nature* **317**, 773 (1985).
3. Allen, J.F. & Misener, A.D. *Proc. R. Soc. A* **172**, 467 (1939).
4. Kapitza, P.L. *J. Phys. Moscow* **5**, 59 (1941).
5. Osheroff, D.D., Richardson, R.C. & Lee, D.M. *Phys. Rev. Lett.* **28**, 885 (1972).
6. Osheroff, D.D., Gully, W.J., Richardson, R.C. & Lee, D.M. *Phys. Rev. Lett.* **29**, 920 (1972).
7. London, F. *Superfluids*. Vol.2 (Dover, New York, 1960).

New ways with crystal growth

Computer simulations of the aggregation process are fashionable because the problems are complicated. But mathematics still has a place.

THAT crystal growth is a complicated process requires no formal proof: how, otherwise, runs the conventional argument, would snowflakes have such a shape? One consequence of considerations such as these is the wealth of numerical simulations which there have been, in the past few years, of how elementary particulates may be added to the surface of a growing aggregate by diffusion from the exterior. This fashion is a reaction to earlier disappointments, and particularly to the recognition that earlier macroscopic models of the growth of aggregates, which distinguished between the various positions on an extending surface only by their macroscopic properties and especially their curvature, have given only poor accounts of what really happens. So it may be a salutary lesson for all concerned that there should now have appeared an analytical calculation of the changing shape of a surface growing by aggregation; people who have been learning how to program increasingly complicated computer machinery may yet find themselves mugging up on differential equations again.

To be fair, the question that has arisen is only part of the whole complicated problem, and may be put like this. In the earliest stages of the growth of a crystal or other aggregate, chance adhesion will be the rule, so that small particles will tend to be rough particles. Given time to reach equilibrium, of course, the roughnesses will become smooth as particles in exceptional positions find less energetically exposed sites at which to settle down, but there will also be circumstances in which a surface growing at a modest rate will become smoother in the course of time as added particles find their way preferentially to the gaps left in the growing surface. To account for the irregularity of snowflakes, either microscopically (by numerical simulation) or macroscopically (by allowing for the extra energy of elements added at places where the curvature is great using the language of surface tension), boils down to supposing that the circumstances are dynamic, not static. The rapidly growing tip of a snowflake crystal is, because of its geometry, exposed to a greater volume of the environment from which further elements may be accreted than would be a re-entrant part of the snowflake surface, while the speed with which newly accreted elements may readjust their positions is so much less than

that with which new elements are added that aberrations are, so to speak, frozen.

The new development is surprisingly straightforward, and the product of the cosmopolitan collaboration of Mehran Kardar of Harvard University, Giorgio Parisi from the University of Rome and Yi-Cheng Zhang from the Brookhaven National Laboratory (*Phys. Rev. Lett.* **56**, 889; 1986). The question they set out to answer is that of how a surface accreting extra elements will grow in the course of time, allowing that the circumstances are midway between those which determine the behaviour of a growing snowflake, where there is no opportunity for newly added particles to readjust their positions, and those that dominate the macroscopic growth of solids which are predisposed towards the emergence of globules of one form or another.

The simplest starting point is the assumption that the normal to a surface will determine the direction in which it will grow locally. There follows a geometrical construction of what happens at the surface, beginning with the assumption that it is possible to measure the distance of the surface as a function of time from some fixed plane; the distance might, for example, be the height of the surface of a growing crystal from the bottom of a vessel in which it is allowed to settle. At any point, the rate of change of the height with time will therefore be determined by the rate of growth normal to the surface and by the square of the gradient of the surface at that point. From this it follows naturally and very simply that the rate of change with time of height from the reference plane is proportional, first, to the curvature of the surface at that point and, second, to the square of the gradient of the surface at that same point. The first term is that which represents the effect of surface tension. The second, which makes the differential equation non-linear, allows for lateral growth. For completeness, the specification of the problem requires that there should be an element of noisiness, represented by a third term, in the gaussian function in time and space, whose only important property is that its average should be zero.

Mathematicians will quickly observe that this non-linear equation describing the evolution of an accreting surface can be converted into a linear form that of a diffusion process. Indeed, in this form the equations are similar to those which des-

cribe the behaviour of branched polymer molecules in which the surrounding environment discriminates against over-extended forms.

The interest of this formulation of the problem of the growth of accreting surfaces is that, in principle, it allows of a global view of what may be happening. If the noisiness of the problem is left out of account, the equations suggest that a one-dimensional growing surface will consist of a sequence of parabolic segments which, as time goes on, will grow at each others' expense, the larger segments swallowing the smaller. If the surface is two-dimensional, there will be patches on the surface which form a comprehensive network covering the whole which again, with the passage of time, allow the larger to swallow the smaller. Specifically, for one-dimensional surfaces, the size of the parabolic segments increases with the $2/3$ power of the time, which is what the numerical simulations suggest.

The noisiness of the problem cannot, in reality, be ignored, but cannot as easily be described in words. But a degree of randomness turns out to be a physical benefit in suggesting analogies with what may happen in other physical circumstances, those of quantum mechanics for example. Indeed, the cleverness of this treatment of the problem of growing surfaces lies in the way it can be transformed from a problem in the solution of differential equations to one much more familiar in quantum electrodynamics, that of describing the time-evolution of a system by the effect of operators on some starting function.

The result, in the surface problem, is a set of simple rules that may describe the simple scaling properties of the problem. That such must emerge follows from the simple observation that, in the case of a one-dimensional surface, a structure which is made up from a sequence of parabolic segments transforms itself in the course of time to another sequence of parabolic segments whose average size has grown by a factor which is some fractional power of the time elapsed. The authors are quick to point out that their model agrees with several of the numerical simulations which have been carried out in circumstances when comparisons are possible. But the importance of this novel treatment is as a physical insight into the aggregation problem when simulations can provide only answers, not understanding.

John Maddox

Model crystals

Penny plain, tuppence coloured

from Robert W. Cahn

A FLOURISHING community of theoreticians seeks to clarify processes in crystals — phase transformations, critical phenomena, radiation damage, melting and vitrification — by simulating them in powerful computers. The latest attempts to model crystals realistically are reported in two papers elsewhere in this issue^{1,2}.

Physical modelling of crystals began in the late 1940s with the two-dimensional soap-bubble raft of Bragg, Nye and Lomer. Given the right bubble radius, such rafts reproduced well the interatomic force law typical for metals, and were used to simulate the motion of dislocations under shear stress. They were also used to model metallic glasses, but for this purpose, two populations of bubbles of different radii were essential to inhibit 'crystallization'³. Such rafts have been well used, notably by Argon⁴, to interpret mechanical properties of metallic glasses.

In the 1960s, Bernal produced his influential liquid model with monosized steel balls kneaded in a football bladder, which was later extended to imitate metallic glass structures. In contrast to the bubble raft, Bernal's model suffered from the fact that the balls were hard and could not imitate the 'soft' interatomic force laws appropriate for alloys: this inadequacy was overcome, both literally and metaphorically, by the introduction of software to allow relaxation of a Bernal model via computer simulation.

More recently, a new family of 'macroparticulate' crystals has appeared in the form of colloidal crystals⁵, regular assemblies of equal submicron-sized spheres, typically of a polymer or of silica. These crystals were not intended as models: their study grew from observations of natural macroparticulate crystals — the crystalline viruses and opal⁶ (which consists of regular arrays of monosized amorphous silica spheres). The first artificial macroparticulate crystals to be made, gem-quality opals, were made by gravity-induced sedimentation of colloidal silica suspensions by Pierre Gilson at his factory in northern France⁷ and in commercial production since 1974. How these opals are stabilized after sedimentation (sintering or low-temperature adhesion) is a commercial secret.

The study of silica (and polymeric) macroparticulate crystals has accelerated rapidly since then (although it is doubtful whether many of the physicists concerned know of Gilson's early triumph). The use of macroparticulate crystal assemblies of ceramic microspheres as 'green' precursors for sintering has now been proposed

and a promising start has been made with silica⁸ (as discussed in these columns⁹).

One of the papers in this issue¹ follows the colloidal approach to crystal simulation: Pusey and van Megen used monosized polymethylmethacrylate spheres, 0.3 μm in diameter and stabilized in suspension by a thin coating of another polymer. The organic suspensions were of high volume fractions, ranging from 0.39 to 0.53. After mechanical homogenization, each suspension was allowed to settle. In the volume fraction range 0.44–0.50, polycrystalline structures have bright opalescence; suspensions below 0.41 do not crystallize at all; whereas in the range 0.41–0.44, a stable liquid-plus-crystal two-phase state forms. But the really novel feature occurs in volume fractions above 0.50, where the viscous suspension remains in a 'glassy' form indefinitely: in effect, this is a kinetically metastable random dispersion which is too dense, and therefore too immobile, to order. Unlike two-dimensional bubble rafts, therefore, in three dimensions it is possible to create a glass from a population of monosized spheres. (Whether this is possible depends on the mobility of the units; be they macroparticles or atoms, pure metals cannot be prevented from crystallizing even at cooling rates in excess of 10^{12} K s⁻¹.) Pusey and van Megen's method could be used to make a form of artificial opal if a way to stabilize the crystals against accidental randomization can be found.

Pusey and van Megen's suspensions differ radically from another kind which also show unexpected properties: Clark and Ackerson^{10,11} demonstrated that very dilute (0.1 per cent) suspensions of highly charged polymer spheres can 'crystallize' into body-centred cubic crystals while remaining suspended, but randomize (melt) after vigorous shear deformation — an observation which could shed light on mechanical instability models of melting for ordinary atomic crystals.

The other paper in this issue, by Georges *et al.*², exploits bubble rafts to simulate, in two dimensions, a microhardness indentation test. A conical indenter is simulated by a suitably shaped (poly)crystalline raft; the test object is a large monosized polycrystalline raft covered by an amorphous 'coating' consisting of a mix of two bubble sizes. The authors do not explain why they chose to test a glass-coated crystalline-simulated sample, but in the light of some recent tests on real glassy wear-resistant coatings, it was a happy choice. The simulated coating proves extremely resistant to penetration, but its

behaviour is intimately affected by its crystalline substrate, which undergoes a largely elastic deformation. On unloading, this elastic distortion recovers and the form of the indentation recovers with it, so much so that the end-result is a small mound instead of an indentation.

This model system with its properties is of considerable interest in relation to recent work on the use of an ultramicrohardness tester, the nanoindenter, invented and named by W.C. Oliver^{12,13}. This instrument is extremely sensitive, using measurements of the vertical displacement of the indenter (as small as 20 nm) to evaluate the residual indentation. (Direct measurement of such minute features would require scanning electron microscopy.) The instrument allows the increase in effective hardness for indents less than 100 nm deep to be confirmed and interpreted¹². It was also found that considerable elastic recovery of the indentation takes place, and that this is essentially caused by a recovery of the 'hinterland' of the indentation, just as in the bubble simulation. But nanoindentations in real crystals do not invert and finish up as mounds.

Oliver and his collaborators have also studied the nanohardness of coated (ion-implanted) metals¹³. Of relevance to the present discussion is the behaviour of steel ion-implanted with both titanium (Ti) and carbon (C). If the concentrations of both are sufficient, the coating is very hard and wear-resistant. Such Ti/C implants are known to be amorphous, and the excellent mechanical properties of the layers are directly correlated with their vitreous nature. Georges *et al.*² agree with these findings in the sense that their amorphous simulated coating proved impenetrable, being always pushed ahead of the moving indenter, but disagree in the sense that the yield stress they observed for the coating was lower than that for the crystalline substrate. Amorphous Ti/C coatings are stronger, harder and more wear-resistant than the underlying steel. □

1. Pusey, P.N. & van Megen, W. *Nature* **320**, 340 (1986).
2. Georges, J.M., Meille, G., Loubet, J.L. & Tolen, A.M. *Nature* **320**, 342 (1986).
3. Simpson, A.W. & Hodgkinson, P. *Nature* **337**, 320 (1972).
4. Argon, A.S. & Kuo, H.Y. *Mater. Sci. Engng.* **39**, 101 (1979).
5. Pieranski, P. *Contemp. Phys.* **24**, 25 (1983).
6. Darragh, P.J., Gaskin, A.J. & Sanders, J.V. *Scient. Am.* **234** (4), 84 (1976).
7. Anderson, B.W. *Gem Testing* 9th edn (Butterworth, London, 1980).
8. Sacks, M.D. & Tseng, T.-Y. *J. Am. Ceram. Soc.* **67**, 526 (1984).
9. Calvert, P.D. *Nature News and Views* **317**, 201 (1985).
10. Clark, N.A., Hurd, A.J. & Ackerson, B.J. *Nature* **281**, 57 (1979).
11. Ackerson, B.J. & Clark, N.A. *Phys. Rev. Lett.* **46**, 123 (1981).
12. Pethica, J.B., Hutchings, R. & Oliver, W.C. *Phil. Mag.* **A48**, 593 (1983).
13. Pethica, J.B., Hutchings, R. & Oliver, W.C. *Nucl. Instrum. Meth.* **209/210**, 995 (1983).

Robert W. Cahn is currently Fairchild Distinguished Scholar in the Keck Engineering Laboratories, California Institute of Technology, Pasadena, California 91125, USA.

Somatic mutation

Do plants evolve differently?

from William J. Sutherland and Andrew R. Watkinson

It has recently been stressed that single plants may consist of a mosaic of genetically different parts^{1,2}. If this is true, then the usual explanations for the operation of natural selection on animals may be insufficient to account for the process of genetic change in many plant populations. The relative importance of somatic and gametic mutations in plants cannot be assessed until the necessary measurements are made. But it is clear that somatic mutation could be important in many plant species.

Modern genetics is based on Weissmann's³ distinction between the cell line and the germ line. Mutations within cells may be expressed but are not passed on to further generations; mutations in the germ line are both expressed and transmitted to descendants. In Weissmann's theory, changes in the germ line are the sole means of evolutionary change. But Weissmann was a zoologist and applied his fundamental principle only to animals. In plants there is no distinction between the germ line and the soma. Growth occurs through the iteration of structural units (modular growth) by one or (usually) many growing points; as a consequence a mutation in a meristem may be expressed in subsequent growth and will then be contained in the pollen or ova. Thus, in plants (and some colonial animals) there is the possibility of genetic variation within an individual and, as a consequence, variation in fitness between modules.

There have been few attempts to measure genetic variability within plants, although there are indications that it may be a widespread phenomenon. Chimaera are common in two species of fern, *Matthia struthiopteris* and *Onoclea sensibilis*⁴; 10 out of 56 of the former and 16 out of 37 of the latter possess mutants. The rates of these mutations are high, with 0.0177 and 0.0341 mutations, respectively, per apical cell per generation. In the herbaceous perennial spring beauty (*Claytonia virginica*) 68 per cent of the population show differences in chromosome number within a plant⁵; of 8,000 plant varieties cultivated in Europe in 1899, 5,000 originated as somatic mutations⁶.

Once genetic differences exist between parts of the same plant there is the opportunity for natural selection to modify the gene frequencies within an individual by the process of differential growth. If the parts possessing the mutation grow faster the mutation may spread; if the mutation prevents successful growth then it will be eliminated — anyone possessing an

ornamental plant with variegated leaves knows that it is essential to remove branches containing normal unvariegated leaves which grow more quickly and would eventually dominate the plant.

The importance of differential growth is also shown in the case of the white clover (*Trifolium repens*) in north Wales. There is considerable genetic diversity within populations⁷ and the growth rate of different genotypes depends on the surrounding vegetation⁸. Those genotypes that grow with the grass *Holcus lanatus* produce more ramets when cultivated among other grasses; other genotypes respond similarly to other species. Thus, the abundance and distribution of different genotypes is determined by differential growth.

In principle it is clear that evolutionary change in a population may originate from somatic mutations. It is essential that somatic mutations should be inherited both sexually and asexually and also that there is differential growth of different genotypes within the same plant. But the importance of somatic mutations is in dispute. Some argue⁹ that the role of somatic mutations is trivial, although the model on which this assertion is based does not take differential growth into account. A later model¹⁰ supports the notion that somatic mutations will produce enough variability to prevent pathogens and herbivores from adapting to all branches on individual host

trees. This model also shows that somatic mutation may be an important source of variation within trees and tree populations. It seems likely that this mechanism is unimportant for species with relatively few meristems (such as peas and maize) but more important for long-lived and extensive species. Thus, individual clones of the aspen (*Populus tremuloides*) can cover 200 acres, be comprised of 47,000 trees and may date back to the Pleistocene¹¹ — it would be incredible if each of these clones was genetically homogeneous.

Although trees provide obvious examples of large, long-lived species, somatic mutation and differential growth could also be important in stoloniferous and rhizomatous plants. Most trees are restricted by a unified architecture that limits the potential for differential growth of branches, whereas in stoloniferous species (such as clover and creeping buttercup) the individual plant fragments as it grows, thus allowing a favourable mutant to spread rapidly. □

1. Whitham, T.A., Williams, A.G. & Robinson, A.M. in *A New Ecology* (eds Price, P.W., Slobodkinoff, C.N. & Gaud, W.S.) 15 (Wiley, New York, 1985).
2. Gill, D.E. & Halverson, T.G. in *Evolutionary Ecology* (ed. Shorrocks, B.) 105 (Blackwell, Oxford, 1984).
3. Weissmann, A. *On Heredity* (Clarendon, Oxford 1883).
4. Klekowski, E.J. Jr *Evolution* 38, 417 (1984).
5. Lewis, W.H., Oliver, R.L. & Luikart, T.K. *Science* 172, 564 (1971).
6. Cramer, P.J.S. *Nature 3 Vers Vol.* 6 (Wetenschappelijke, Haarlem, The Netherlands, 1907).
7. Cahn, M.A. & Harper, J.L. *Heredity* 37, 309, (1976).
8. Turkington, R. & Harper, J.L. *J. Ecol.* 67, 245 (1979).
9. Slatkin, M. in *Evolution in Honour of John Maynard Smith* (eds Greenwood, P.J., Harvey, P.H. & Slatkin, M.) 19 (Cambridge University Press, 1984).
10. Antolin, M.F. & Strobeck, C. *Am. Nat.* 126, 52 (1985).
11. Kemperman, J.A. & Barnes, B.V. *Can. J. Bot.* 54, 2603 (1976).

William J. Sutherland and Andrew R. Watkinson are at the School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

Palaeoclimatology

Bears versus beetles

from Peter D. Moore

THE INTERNATIONAL Geological Correlation Programme recently stated the view that there were two separate warm periods during the last (Ipswichian) interglacial period, and that they can be defined on the basis of their mammalian fauna¹. The final (Devensian) glaciation is generally believed to have reached its maximum geographical extent 18,000 years ago (18,000 BP) and by 14,000 years ago Britain was essentially deglaciated. But there are many problems in elucidating the climatic sequence of the previous 200,000 years.

Fossil mammalian bones are an important source of evidence in Quaternary palaeoecology, but their use in climatic reconstruction raises some problems. Such material can be identified with con-

siderable accuracy but compared with other types of fossil evidence, such as pollen, diatoms and beetles, mammalian bones are rare. This means that sufficiently large populations of bones to allow statistical analysis at any site are infrequent, whereas smaller and more abundant fossils, such as pollen grains, lend themselves to this approach. Climatic conclusions from mammalian fossils are also complicated by the relatively large range of tolerance of mammals compared with cold-blooded invertebrates. Thus, one can argue that beetles are better climatic indicators than bears. But some aspects of climatic reconstruction in the late Quaternary are now under debate as a result of finds of mammals bones in Britain.

Until recently it was thought that the

Ipswichian was the only interglacial during the 200,000 years preceding the Devensian, but there are increasing difficulties in reconciling the evidence from various sites presumed to be of this age¹. Fossil assemblages in Britain at sites such as Bobbitshole, Suffolk² and Trafalgar Square, London³ contained a typically warm climate fauna, including hippopotamus, and lacked any steppic species, such as horse.

At other British sites, such as Marsworth, Buckingham and Stanton Harcourt, Oxford, deposits with a similar pollen spectrum and therefore assumed to be Ipswichian yield fossils such as mammoth and horse, indicative of colder conditions. There is evidence that these mammoth-bearing strata are separated from those containing warm, hippopotamus fauna by a very cold intervening phase. The mammoth/horse episode is the earlier of the two and can be correlated with oxygen isotope stage 7. The hippopotamus event has been dated by uranium and thorium isotopes to between 135,000 and 114,000 BP at Victoria cave in Yorkshire⁵, which corresponds to oxygen isotope substage 5e.

Precisely what happened between the time when hippopotami were wallowing in the British wetlands and the intense glaciation of 18,000 years ago is still uncertain. There is evidence of a warm climate about 60,000 years ago at Chelford, Cheshire, when pine, birch and spruce were the major dominants of the pollen flora⁶. But the radiocarbon dating of this site is not entirely satisfactory as it extends beyond the timescale in which the method is reliable.

Pollen data showing the increasing cold nature of the flora following the Ipswichian Interglacial are available from a site at Wing, Leicester⁷, and a marine pollen sequence has been produced from Fjøsanger Fjord in western Norway⁸. The latter site shows an abrupt decline in forest vegetation and a spread of grassland as the last glacial episode began.

Further evidence concerning the early part of the last glacial has now been obtained from another cave deposit in Yorkshire, only 23 km away from the Victoria cave site. The Stump Cross cave⁹ has been found to contain bone remains of wolf, reindeer and wolverine, which constitute a distinctly tundra or boreal mammal assemblage. In many respects the assemblage resembles that found in the mammoth-bone houses of the Ukraine¹⁰, dating from the end of the last glaciation, although no mammoth bones are found at Stump Cross. The Stump Cross bones are associated with calcite deposits that have been subjected to uranium-series radiometric dating and provide dates ranging between 84,000 and 81,000 BP which, if taken at their face value, places the bone collection in the younger part of

oxygen isotope stage 5, between the hippopotamus fauna of Victoria cave and the spruce forest of the Chelford Interstadial.

There must have been a considerable fall in temperature between 120,000 and 85,000 BP to account for this change of fauna, but closer interpretation in climatic terms is not simple. The three major animal species were probably inter-related as predators and prey, the wolf and the wolverine preying on reindeer. But damage to the bones by splintering and chewing suggest that the bone hordes were accumulated by wolverines that also preyed on wolves and even on one another. The physical and biological setting of this drama is difficult to reconstruct precisely because wolverines and the other species represented are found not only in open tundra areas, but also in the coniferous forest (taiga) zone. So, on the basis of evidence from mammalian bones, it is still not possible to determine whether the earliest phase of the Devensian was colder than the conditions in the Chelford Interstadial. It does not provide an answer to the origin of the till underlying the Chelford deposits, which may belong to an earlier glaciation¹¹. Climatic deterioration since the day of the hippopotamus had clearly

been substantial, but was it sufficient to result in a pre-Chelford glaciation in the British Devensian?

The story has now become even more complicated by a report earlier this month from Bacon Hole, Wales¹² of sediments dated at 81,000 BP that contain a clearly warmth-demanding fauna including straight-tusked elephant (*Palaeoloxodon*), rhino (*Dicerorhinus*) and hyaena (*Crocota*). This will certainly cast the interglacial cats among the periglacial pigeons. Perhaps the beetles will still have the last word. □

1. Shotton, F.W. *Quat. Newslett.* **45**, 28 (1985).
2. IGCP Project 24 *Quat. Newslett.* **39**, 20 (1983).
3. West, R.G. *Phil. Trans. R. Soc. B241*, 1 (1957).
4. Franks, J.W. *New Phytol.* **59**, 145 (1960).
5. Gascoyne, M., Currant, A.P. & Lord, T.C. *Nature* **294**, 652 (1981).
6. Simpson, I.M. & West, R.G. *New Phytol.* **57**, 239 (1958).
7. Hall, A.R. *Phil. Trans. R. Soc. B289*, 135 (1980).
8. Mangerud, J., Sønstegeard, E. & Sejrup, H.-P. *Nature* **277**, 189 (1979).
9. Sutcliffe, A.J. *et al. Quat. Res.* **24**, 73 (1985).
10. Müller-Beck, H. in *Paleoecology of Beringia* (eds Hopkins, D.M., Matthews, J.V., Schweger, C.E. & Young, S.B.) 329 (Academic, London, 1982).
11. West, R.G. *Pleistocene Geology and Biology* 2nd Edn 295 (Longman, London, 1977).
12. Stringer, C.B., Currant, A.P., Schwartz, H.P. & Collcutt, S.N. *Nature* **320**, 59 (1986).

Peter D. Moore is Reader in Ecology at the Department of Biology, King's College (KQC), 68 Half Moon Lane, London SE24 9JF, UK.

Igneous petrology

Archaean mantle models

from E. G. Nisbet

ONE of the chief interests of igneous petrology lies in establishing the nature and the history of the Earth's mantle, the region below the differentiated crust and above the central core. We know very little about its composition and structure, how the continents become differentiated from it and how it has evolved thermally. In the past few years new data have allowed petrologists to produce an exotic array of speculative models of the state of the mantle in Archaean time (before 2.5×10^9 years ago). Many of these models are based on inferences drawn from komatiites (magnesian lavas that erupted at temperatures up to 1,650°C). Recently, Herzberg and O'Hara¹ and Walker² have used experimental data on the melting behaviour of natural samples to suggest that the upper mantle (shallower than about 670 km) has a very unlikely composition which is best explained if it originated as a melt, leaving the lower mantle as residue.

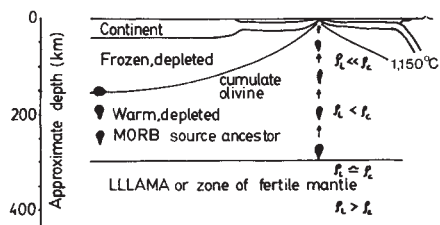
Recent experimental work^{3,4} on 'fertile' nodules from great depths in South African diamond pipes^{5,7} shows that the solidus and liquidus temperatures (the temperatures at onset and completion of melting) converge at pressures of 100–150 kbar (equivalent to depths of 300–450 km). If the fertile nodules are representative of mantle composition, then the

close approach of solidus and liquidus has major implications. Walker² has shown that such an approach to convergence is most improbable if the upper mantle is a random aggregate of crystals collected during accretion or as a residue from melting. More likely is that the upper mantle samples show convergence of solidus and liquidus because they were themselves once liquid compositions produced by high-pressure congruent crystal-liquid interaction. If the samples are indeed representative of the upper mantle, the implication is that the upper mantle was itself a melt, derived either by the accumulation of melts from the lower mantle, as suggested by Herzberg and O'Hara¹, or from a vast globe-encircling magma ocean⁶.

How definite are these conclusions? There are, of course, experimental uncertainties: errors in the measurement of temperature are probably fairly large, but not large enough to undermine the convergence of solidus and liquidus. More important is how representative the nodule samples are of the upper mantle. Diamonds are known to be very old, and have been held in the cool continental lithosphere for up to 3×10^9 years at depths of about 150 km. Possibly the fertile nodules were kept there with the diamonds, and may thus represent liquids

that ascended and were trapped by being frozen into the base of the continents in mid-Archaean time, since being reworked by metasomatism and the accidents of eruption. If so, the nodules may not be particularly good guides to bulk mantle compositions.

There are other lines of evidence that help to elucidate the nature of the upper mantle. Komatiites erupted in many locations on the Earth's surface in Archaean time, and must have been derived from a mantle source. It is possible that they come from a source similar to the fertile nodules, and it has also been suggested³ that they were produced by very small degrees of partial melt of material similar to the nodules, at pressures in excess of 70 kbar. But, the near-chondritic composition of the trace elements in the lavas suggest instead that they originated as fairly high degrees of melt of a mantle parent. One possible model of the early Archaean mantle⁶ is based on the prediction that at great depth komatiite liquids are denser than the coexisting olivine. If this is the case, it is possible that large



Model of the possible structure of the Archaean mantle. All depths very poorly constrained. Blobs, rising melt; ρ_L , density of liquid; ρ_c , density of crystals; Some blobs erupt as komatiite or derived liquids with varying degrees of contamination, others are trapped and frozen into the continental lithosphere.

laterally linked magma accumulations (LLAMAs) existed in the Archaean Earth, trapped by chemistry and density as a magma ocean whose top surface was at a depth of about 250–300 km (depending on the level of the supposed density crossover). Such a magma ocean would act as a source for komatiitic liquids.

Isotope evidence also provides an interesting picture of the Archaean upper mantle. The mafic lava sequences typically seem to have come from sources which had positive ϵ_{Nd} values, something that is normally thought to imply that the source region had previously undergone extraction of a magma enriched in incompatible elements. Arndt (personal communication) has pointed out that a magma ocean source would probably be little depleted and thus is an implausible direct source of the parent liquid for Archaean mafic lavas. Furthermore, there is considerable variation in the compositions of komatiites, which is difficult to reconcile with the notion that they all have a common source.

Very little is yet known about the com-

position of the early upper mantle. If the experimental data from nodules is indeed a good guide, then the mantle may have been extensively molten or perhaps a serial accumulation of melts. But the nodules may not be representative. Similarly, the komatiites may imply the presence of a magma ocean, but is this consistent with their chemical variation and the isotopic evidence for a depleted source?

A speculative reconstruction is shown in the figure: possibly the mantle was stratified, with a depleted light layer⁵ over a magma ocean. Komatiites may have

been derived in various ways from the ocean or from the depleted layer, or could have been isotopically and chemically contaminated on ascent. □

1. Herzberg, C.T. & O'Hara, M.J. *Geophys. Res. Lett.* **12**, 541 (1985).
2. Walker, D. *Contrib. Mineral. Petrol.* (in the press).
3. Takahashi, E. & Scarfe, C.M. *Nature* **315**, 566 (1985).
4. Scarfe, C.M. & Takahashi, E. *EOS* **66**, 1130 (1985).
5. O'Hara, M.J., Saunders, M.J. & Mercy, E.L.P. *Phys. Chem. Earth* **9**, 571 (1975).
6. Nisbet, E.G. & Walker, D. *Earth planet. Sci. Lett.* **60**, 105 (1982).
7. Richardson, S.M., Erlank, A.J. & Hart, S.R. *Earth planet. Sci. Lett.* **75**, 116 (1985).

E.G. Nisbet is in the Department of Geological Sciences, University of Saskatchewan, Saskatoon S7N 0W0, Canada.

Tropical plant ecophysiology

Research in South-East Asia

from Barry Osmond

HUMAN activities in the tropics are responsible for large-scale and rapid ecological changes. In South-East Asia the clearing of forests threatens to eliminate or seriously reduce the populations of many commercially and scientifically valuable plant species within the next two decades. These threatened species must be maintained in managed ecosystems which contribute to the economic well-being of the region, and not solely in reserves dedicated to the conservation of genetic resources. Whole new ecological systems, based on agroforestry and mixed cropping, are being constructed (Torquebiau, E.F. *Agroforestry Systems* **2**, 103; 1984). All of these activities are taking place in a near-vacuum so far as understanding of the functional relationships within tropical plant communities is concerned. A recent international conference in Bogor, Indonesia* focused on the needs and opportunities for plant ecophysiological research in South-East Asia, and established a regional network of ecologists, agronomists and foresters to promote these studies.

Two main themes emerged. First, the understanding of temporary environmental change and the biological response to this change in tree seedlings following gap formation is primarily relevant to forest management (A. Mohammed, Forest Research Institute, Kepong). Second, understanding the problems of plant stress physiology in permanently changed, often degraded, grassland and cultivated land is relevant to the productivity of indigenous and introduced crops, most of which perform poorly under tropical conditions (F. Muhadjir, Research Institute for Food Crops, Bogor).

Many of the foundations of tropical

plant ecophysiology were laid in South-East Asia by early botanical explorers such as Schimper in the late nineteenth century, and by pioneering experimentalists like Stocker who worked in the Bogor botanical gardens in the 1920s. Unlike the American tropics, investigation of plant biology in South-East Asia was in the hands of botanists, rather than naturalists, and there is now a large body of ecological data for the region (for example, Whitmore, T.C. *Tropical Rainforests of the Far East* 2nd edn; Clarendon, Oxford, 1984). However, the rate of exploitation has far outstripped our understanding of functional relationships in tropical vegetation, all of which is summarized in one slim volume (Medina, E., Mooney, H.A. & Vazquez-Yanes, C. (eds) *Physiological Ecology of Plants of the Wet Tropics* Junk, The Hague, 1984). Our ignorance of vegetation dynamics and the principles of plant/environment interactions in tropical systems has led to coexistence of contradictory theory and practice — to management based on faith rather than fact.

Many tropical plant biologists now believe that understanding of forest systems hinges on vegetation responses to gap formation, called patch dynamics (Watt, A.S. *J. Ecol.* **35**, 1; 1947). It is now known that leaves of trees grown in the shade possess an astonishing ability to use light flecks of 30–60 seconds' duration which depends both on prior light experience and on significant post-illumination CO_2 fixation. These induction and capacitance components of light-fleck use (R.W. Pearcy, University of California at Davis) are drastically modified during the sustained increase of light which follows gap formation. It was previously thought that plants of deeply shaded habitats were excluded from open habitats by photo-destruction of the photosynthetic apparatus (photoinhibition), and their inability to acclimatize to high light. However, it

*SEAMEO-BIOTROP South-East Asian Regional Center for Tropical Botany, 4–6 December 1985, supported by IDRC, Canada; ICSU International Biosciences Networks; and UNESCO, and organized by Dr Wongchan Wongkaew, Botany Department, Kasetsart University, Bangkok Bangkok.

now seems that tree seedlings on the forest floor, and understorey shade plants such as *Alocasia*, persist in bright light by maintaining a dynamic compromise between limited acclimation and chronic photoinhibition.

New, simplified techniques, which bring the diagnosis of photosynthetic acclimation and photoinhibition within easy reach, will permit ecophysiological analyses of these factors in vegetation dynamics. For example, disk O_2 electrodes originally developed for studies of photosynthetic induction at the Research Institute for Photosynthesis, University of Sheffield, have been adapted to measure quantum yield and room-temperature fluorescence (Walker, D.A. & Osmond, C.B. *Proc. R. Soc. B*, in the press). In training courses before the conference, the electrodes were applied to assess the lack of acclimation and extent of photoinhibition in ratan, an important forest product of South-East Asia which is very fastidious with regard to light. The same system has been used to probe the light responses of photosynthesis in tropical forest epiphytes, which may have a role in branch fall and gap formation. Many of these epiphytes carry out CO_2 fixation via crassulacean acid metabolism (more commonly associated with succulent plants in high-light, desert habitats), which, with intrinsically high quantum yields, appears to confer photosynthetic advantages under conditions of low light and high temperature (Winter, K., Osmond, C.B. & Hubick, K.T. *Oecologia* 68, 224; 1986).

Integration of these insights into the performance of tropical plants with key ecological observations, such as leaf phenology pattern (R. Corlett, National University of Singapore) will help to define the productive structure of natural, disturbed and managed forest systems. The same ecophysiological approaches are needed to understand the role of environmental stresses, such as high temperature, inadequate nutrition and water, in limiting the productivity of agriculture of cleared land in the tropics. Even crops of tropical origin, such as maize and sorghum, perform poorly compared with temperate regions, whereas their wild counterparts such as *Imperata* are very aggressive. It is increasingly clear, for example, that high light intensities exaggerate the effect of stress on the photosynthetic apparatus. The light-avoiding movements of leaves of the tropical legume *Siratro*, which mitigate water and high temperature stress (Ludlow, M.M. & Björkman, O. *Planta* 161, 508; 1984), is but one example of the ecophysiological insights that could provide a guideline for improvement of indigenous and introduced crops in the tropics. □

Barry Osmond is Professor of Environmental Biology at the Australian National University, Box 475, Canberra City, ACT 2601, Australia.

Creationism

On the tracks of men and money

from Tony Thulborn

IN their attempts to discredit the theory of evolution, creationists have made one claim that captures the public's imagination like no other — the claim that humans coexisted with dinosaurs. The evidence for this notion is said to reside in the Cretaceous limestones (about 120 million years old) of the Paluxy River, Texas, where markings alleged to be human footprints are preserved alongside the tracks of dinosaurs. These dubious 'man-tracks' figure prominently in creationist literature and have even been featured in several films, including the widely distributed *Footprints in Stone*. There is no doubt that the tale of the man-tracks is one of the most popular draw-cards held by the creationists. Consequently it is surprising to learn that the tale has now been publicly renounced by a leading American creationist.

On 7 January, John D. Morris, author of a standard creationist text on the man-tracks (*Tracking those Incredible Dinosaurs and the People who Knew Them* CLP, San Diego, 1980) made a startling confession: in a lecture at the University of New South Wales, he admitted that he could find no scientifically acceptable evidence of human tracks in the Paluxy limestones and agreed that it would be improper for creationists to continue citing man-tracks as evidence against evolution. Morris repeated his conclusions in a leaflet published by the Institute for Creation Research (*Impact* 151, 1; 1986) and has also announced that *Footprints in Stone* will be withdrawn.

These welcome retractions came during the course of the 1986 Creation Science Summer Institute, a creationist symposium organized by Australia's Creation Science Foundation (CSF). Why did Morris change his mind? Apparently many of the man-tracks in the Paluxy River have now been weathered to such a degree that their true nature is becoming embarrassingly obvious: the vaguely elliptical markings heralded as human footprints have gradually weathered into the unmistakable outlines of three-toed dinosaur tracks. In the course of his lecture Morris expressed astonishment that the clarity of fossil footprints could improve, rather than deteriorate, as a consequence of natural weathering. This process is well known by professional palaeontologists but was unfamiliar to Morris, who suspected that fanatical anti-creationists had been defacing the evidence.

Morris chose to make his retraction at a crucial moment in the history of Australia's creationist movement. His forthright

confession may go some way towards restoring the public image of CSF, which has suffered an onslaught of damaging criticism in past weeks. The most destructive criticisms appeared in a book (*Creationism, an Australian Perspective* Australian Skeptics, Melbourne, 1986) that was published only 72 hours before CSF began its summer institute. The editors of the book, Martin Bridgstock (Griffith University) and Ken Smith (Queensland University), present a series of short essays explaining the scientific and intellectual bankruptcy of CSF's arguments. More unusually, they also delve into the organization's financial dealings, based on a search of the Corporate Affairs Office in Brisbane.

The financial statements of CSF for the years 1983–84 and 1984–85 contain some intriguing facts and figures. It appears that CSF disposed of more than \$A92,000 under the heading of an 'extraordinary item' (sic). CSF still declines to give a detailed explanation of where the money went. Next, it transpires that CSF does not spend much money on scientific research. The statement for the year ending 31 March 1983 lists an outlay on research of \$A398; in the following year research expenditure soared to \$A860, less than 0.2 per cent of the organization's half-million dollar income.

Most remarkable of all is the discovery that CSF received federal government funding to the amount of \$A24,698 during the two-year period. The fact that federal funds were delivered to CSF in the form of an 'export market development grant' raises an interesting question: what does the organization intend to export? The obvious answer is that it plans to export religious materials, which is evident from the fact that CSF officially declares its principal business to be: "The preaching of the Gospel, Christian Education, the accumulation and presentation of scientific evidence relevant to Creation, publication of a magazine and the sale of books, tapes and filmstrips".

If it can be demonstrated that the exports of CSF are religious in nature it is likely that its federal funding will be ruled unconstitutional. An article in the *Melbourne Age* (4 January 1986) reports that lawyers are now investigating the possibility of a challenge under section 116 of the Australian constitution, which expressly prohibits the government from establishing a religion. □

Tony Thulborn is in the Department of Zoology, University of Queensland, St Lucia, Brisbane, Australia 4067.

Colloid chemistry

Applications of microemulsions

from B.H. Robinson

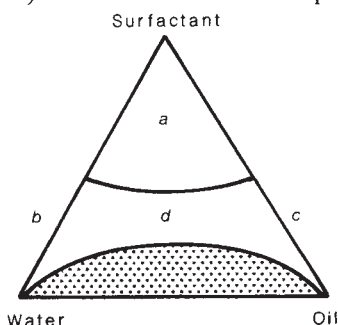
It is well known that oil and water do not mix. However, if a chemical dispersing agent, known variously as a surface-active agent, surfactant or detergent, is added then a 'microemulsion' liquid dispersion may well form. The surfactant, located at the interface between the two phases, dramatically lowers the oil-water interfacial tension. The dispersions are optically transparent and are generally thought to be thermodynamically stable. Until recently, physicists and physical chemists interested in the properties of the liquid state did not regard such systems worth serious study, believing that they were far too complex and irreproducible in their behaviour. But this situation is now changing rapidly and microemulsions are attracting wide interest in both industrial and academic research laboratories. On page 338 of this issue¹, and in a preceding paper², J. Tabony reports the results of a structural study of a surfactant-stabilized dispersion containing equal volumes of oil and water that indicates the presence of well-defined structural domains.

Most experimentalists have avoided the composition region in the middle of the triangular phase diagram (see figure), favouring the simpler 'precursor' systems along the water-oil and oil-surfactant sides of the triangle. In the absence of oil, surfactants spontaneously self-organize in water at low concentrations to form micelles, and as the surfactant concentration is increased, a variety of lyotropic liquid crystalline structures are formed. These soap-water systems are of great importance in detergency and are now quite well understood. In a similar way, certain surfactants, including phospholipids, form inverted micelles, which can readily take up water to form microstructural domains in which water is dispersed in oil in the form of essentially monodisperse aqueous droplets. Their size can be precisely controlled by varying the amounts of water and surfactant. In contrast, previous studies in the composition region (*d* of the figure) suggest the existence of open 'bicontinuous' sponge-like structures in which the surfactant forms an interface of rapidly fluctuating curvature but in which the net curvature (time- or space-averaged) is near zero³.

The results reported by Tabony in this issue were obtained using one of a range of neutron-scattering instruments which are available for structural and dynamic studies at the Institute Laue Langevin, Grenoble. Much new information on size, structure and motion of a whole range of colloidal dispersions is now becoming

available as a result of the successful exploitation of the Grenoble facility, and research will continue also at the pulsed-neutron source in the UK.

Neutrons are particularly useful for studies on liquid dispersions because the method of contrast variation (selective deuterium labelling of a component of the system) can reveal the internal structure of an individual droplet and motions of particular components. The accessible wavelength range is such that inter-droplet interactions are also readily studied using small-angle scattering methods⁴. Molecular motions on a very short (nanosecond) timescale can also be probed



Schematic triangular phase diagram of an oil-water-surfactant system. Shaded region, water-oil immiscibility region (2-phase); *a*, liquid-crystalline phases; *b*, micelles in aqueous solution; *c*, reversed-micelles in oil; *d*, concentrated single-phase microemulsion domain.

using inelastic scattering techniques.

Many detailed triangular phase diagrams for a range of surfactants are already available (for example, ref. 5) and furthermore, water can be substituted by several other polar solvents, for example, glycerol and formamide. Theoretical papers have stressed the importance of the shape of the surfactant molecule (and its natural curvature when adsorbed at the oil-water interface) in influencing the structures which form⁶, so we can now predict structures more effectively.

Why is industry interested in such microemulsions? There are a number of reasons. For example, the oil industry is keen in the longer term to exploit surfactant-enhanced tertiary oil recovery methods for which rheological properties of the microemulsion are important, and to explore the possibilities of gasoline substitutes based on liquid blends with alcohol. There are obvious applications in the food and cosmetics industries, particularly using phospholipid surfactants. Single component thermotropic (low molar mass) liquid crystals have already found wide application as electro-optic

display devices in watches and calculators. Applications of the lyotropic analogues, based on surfactant-water systems, have not been so forthcoming but the extra dimension introduced by the oil component produces a rich diversity of structures with potentially interesting rheological and electrical properties. Characterization of structure and interactions at the molecular level is the first step towards the exploitation of such new materials.

The possibility of precisely controlling the size and stability of the microstructure domains suggests interesting applications for liquid-membrane technology, and the compartmentalized liquid structures of high surface area provided by the oil-water dispersions suggest their use as a novel and versatile medium for chemical synthesis. The potential in this field is indicated in a recent review on enzyme processes in microemulsions⁷. Enzymes retain their activity in such dispersions and in an oil-rich medium the thermodynamic equilibrium of reactions is shifted in such a way that the possibility exists for catalysing reactions in the reverse direction to that favoured in aqueous solution, a process known as reverse enzyme synthesis. In this way, esters can be synthesized rapidly under mild conditions from carboxylic acids and alcohols, or oligopeptides can be prepared from amino acids.

Another potential use of the fluid but structured microdomains of the type reported by Tabony is in the preparation of microscopic particles of a desired size or shape reflecting the structure within the liquid dispersion. Small spherical metallic particles of, for example, platinum⁸ and cobalt boride have already been prepared successfully in water-oil microemulsion systems using this approach and the preparation of particles of other shapes should be possible in microemulsions which have a long-lived non-spherical microstructure. Such particles show catalytic activity in selective hydrogenation reactions⁹. J.H. Fendler has stressed the wide range of chemical processes, such as artificial photosynthesis, that may be facilitated by reaction at interfaces¹⁰. Microemulsions and the so-called 'liquid crystal microemulsions' as described by Tabony look to have considerable potential for exploitation in the near future. □

1. Tabony, J. *Nature* **320**, 338 (1986).
2. Tabony, J. *Nature* **319**, 400 (1986).
3. Gennes, P.G. & Taupin, C. *J. phys. Chem.* **86**, 2294 (1982).
4. Ottewill, R.H. (in Goodwin, J.W. ed. *Colloidal Dispersions* R.S.C. Spec. Pub. 43, 143; 1982).
5. Ekwall, P. *Adv. Liquid Cryst.* **1**, 1 (1975).
6. Israelachvili, J.N., Mitchell, J.D. & Ninham, B.W. *JCS Faraday Trans.* **72**(2), 1525 (1976).
7. Luisi, P.L. *Ang. Chemie, Int. Ed. Engl.* **24**, 439 (1985).
8. Boutonet, M., Kizling, M., Stenius, P. & Mair, G. *Colloids Surf.* **5**, 209 (1982).
9. Lufimpadio, N. *et al. Surfactants in Solution* (eds Mittel, K.T. & Lindman, B.) Vol 3, 1483 (Plenum, New York, 1983).
10. Fendler, J.H. *Chem. Engng. News* 2 January, 25 (1984).

B.H. Robinson is in the Chemical Laboratory, University of Kent, Canterbury CT2 2NH, UK.

Visual cortex

What layer 6 tells layer 4

from Simon LeVay

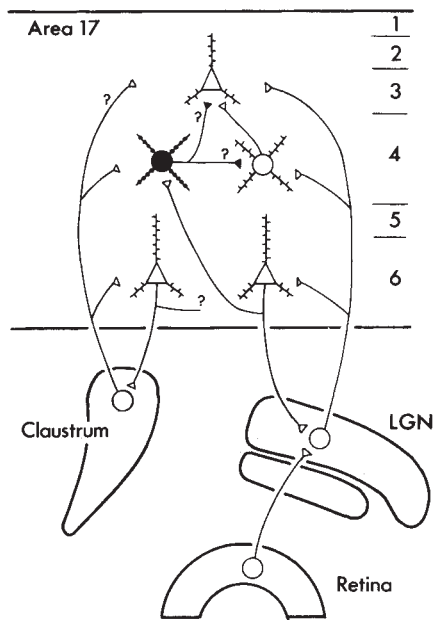
THE primary visual cortex (area 17) of the cat has been the object of intense study since Hubel and Wiesel's pioneering work, first published 24 years ago (*J. Physiol., Lond.* **160**, 106; 1962). As a result, we now know the characteristic shapes and connections of many of the major cell types, have categorized the visual responses of neurones in area 17 and matched morphological and functional cell types using intracellular recording and staining techniques. Taken together, these data allow increasingly precise hypotheses about the generation of specific response properties of cortical neurones. Elsewhere in this issue, J. Bolz and D.C. Gilbert put forward one such hypothesis and provide experimental support for it (*Nature* **320**, 362; 1986).

Most neurones in area 17 are orientation-selective, that is, a neurone responds best to an elongated bar of light at a particular orientation that is swept broadside across its receptive field. Many cells have specific requirements with respect to the length of the bar (its dimension along the orientation axis). Some cells respond increasingly well as the stimulus length is increased up to some optimal value (typically 1–3 degrees of visual angle), but respond less well, or cease to respond completely, if the stimulus is lengthened further. The output of these 'end-inhibited' cells thus conveys information about both the orientation and the length of the stimulus. End-inhibited cells are common in the upper cellular layers of area 17 (layers 2, 3 and 4) but are rare in layer 6, the deepest layer. Layer 6 contains many cells which respond poorly to short bars; instead they respond increasingly well as the length of the bar is increased to quite high values (4–16 degrees). Bolz and Gilbert suggest that the activity of these cells is responsible for the end-inhibition of cells in the upper layers.

The circuitry by which this could be achieved is illustrated in the figure. Many pyramidal neurones in layer 6 have axonal branches that ascend to layer 4, where they terminate predominantly on the varicose dendrites of spinefree stellate cells. (It is known that the output of the spiny dendrite stellate cells is largely to the pyramidal neurones of layers 2 and 3.)

According to Bolz and Gilbert's hypothesis, presentation of a short bar (say 2 degrees long) activates layer 4 cells but has little effect on layer 6 cells, whereas pre-

sensation of a long bar (say 8 degrees) activates layer 6 cells, which excite the spinefree cells of layer 4 and suppress the response of the spiny dendrite cells, producing end-inhibition. For this scheme to work the connections must be specific to sets of neurones sharing a common orientation preference — this is taken care of by the fact that interlaminar connections are predominantly vertical and cells sharing a common orientation preference are organized into vertical columns.



Some intrinsic and subcortical connections of area 17 that may be involved in the generation of end-inhibition. Retinal neurones drive neurones in the lateral geniculate nucleus (LGN), which in turn provide input to neurones in cortical layers 3, 4 and 6. Many layer 6 cells have ascending branches which contact spinefree stellate cells (shown in black) in layer 4. These cells are thought to inhibit the spiny stellate cells of the same layer and/or the pyramidal cells of layer 3. Descending axons from layer 6 neurones innervate the LGN and the claustrum. Claustral neurones send a return projection to area 17 that ends in all layers, but most heavily in layer 4. All the cortical neurones drawn here lie in the same orientation column. Uncertain connections are indicated by the question marks.

To test their model, Bolz and Gilbert recorded from cells in the upper cortical layers while reversibly inactivating a patch of layer 6 cells immediately below by direct injection of the inhibitory neurotransmitter γ -aminobutyric acid. During the block, which lasted for a few minutes after each injection, most upper-layer cells lost all or part of their end-inhibition — that is, they responded to long as well

as to short bars. Other properties of the neurones (orientation and direction selectivity) should be largely unaffected.

In terms of the history of ideas about the laminar organization of the cortex, Bolz and Gilbert's results suggest something of a compromise. Earlier workers stressed the notion of reverberating interlaminar circuits or of sequential processing of visual signals from layer to layer. According to these ideas, inactivation of a single layer should largely or totally disable cortical function. In contrast, more recent observations suggest a surprising degree of functional independence of the layers. For example, layer 4, long thought to be the major input layer of area 17, can be silenced without seeming to affect responses in layers 2 and 3 (Malpeli, J.G. *J. Neurophysiol.* **49**, 595; 1983). Bolz and Gilbert show that inactivating a single layer can eliminate a single aspect of cortical function while leaving other aspects intact, and their results illustrate how columnar organization, at least in the orientation domain, simplifies the elaboration of response properties.

What is the specific wiring responsible for these processes? As well as sending axons to layer 4, layer 6 neurones also provide the descending arms of two loops that connect area 17 with subcortical nuclei, the lateral geniculate nucleus and the claustrum (see figure), therefore these descending pathways are also silenced by the γ -aminobutyric acid injections. The cortico-claustral loop has previously been implicated in the generation of end-inhibition in area 17 (Sherk, H. & LeVay, S. *J. Neurosci.* **3**, 2121; 1983), although its contribution is only partial. The role of the massive descending pathway to the lateral geniculate nucleus is something of a mystery. But geniculate neurones have an antagonistic centre-surround type of receptive-field organization, which gives them some degree of end-inhibition when they are stimulated with elongated bars. There are hints that the cortico-geniculate pathway could modulate this effect (Tsumoto, T. *Expl Brain Res.* **32**, 345; 1978) thus providing another route by which layer 6 cells could indirectly influence end-inhibition in the upper cortical layers. More precise experiments are required to elucidate the contributions of these pathways to end-inhibition.

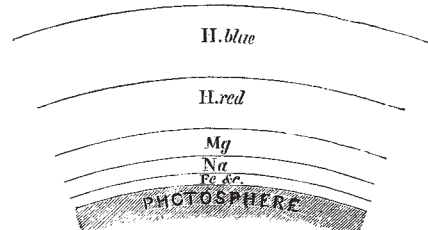
To what extent can Bolz and Gilbert's findings be extended to other cortical areas or to other species? The almost total lack of data allows free rein for speculation. Perhaps the general role of layer 6 is to determine whether increased local cortical activity is matched by heightened activity of neighbouring regions, and, if so, to suppress the activity of upper layer cells. This would be equivalent to the lateral inhibition seen at early levels of

sensory processing (for example, the centre-surround organization of retinal ganglion cells), but it would affect whatever parameters are laid out along the surface coordinates of a particular cortical area. This kind of hypothesis can be tested when the analytical approaches now being applied successfully in area 17 are extended to other parts of the cerebral cortex. □

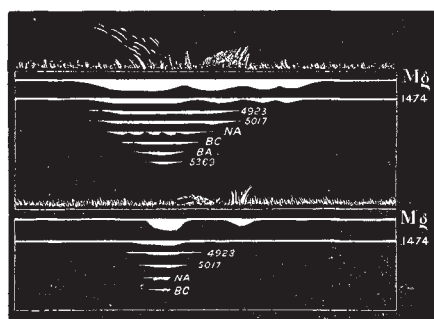
Simon LeVay is at The Salk Institute, PO Box 85800, San Diego, California 92138-9216, USA.

100 years ago THE SUN AND STARS The Chromosphere

In what has gone before we have been chiefly occupied with a discussion of the various chemical materials which we can trace in those cavities in the photosphere which we call spots. We have now to begin the consideration of the chemical materials which can be traced in that solar envelope which lies immediately over the photosphere, I mean the chromosphere: so that eventually we may endeavour to make a comparison between the chemical materials in the spots and in the chromosphere, which are supposed to lie, and which in fact really do lie, at about the same height in the solar atmosphere, with, however, the enormous difference that we know the spots are caused by the descent of materials coming down from above, and we do not know at present that that is true with regard to the substances in the chromosphere.



Early hypothesis of the arrangement of materials in the Sun's atmosphere. H = hydrogen; Mg = Magnesium; Na = sodium; Fe, & c. = iron and the other elements of high atomic weight.



Welling up of vapours

Now, the chromosphere we will take roughly, as it varies in height from year to year, and from latitude to latitude, to be between 5000 and 10,000 miles high. It is not only bright at the bottom — so bright, very often, that in eclipses, when the bottom is seen, observers imagine that the sun has reappeared — but it is exquisitely coloured at the top, and colours very often being scarlet, crimson, green, yellow, and so on. As ordinarily observed, the simple chromosphere varies very considerably.

From *Nature* 33 499, 25 March 1886.

Low-luminosity stars

How no(w) brown dwarfs?

from Virginia Trimble

WHOLE careers and industries are built on extremes — the highest jump, the oldest fossil, the fastest horse. Thus it is perhaps not surprising that astronomers should, now and again, ask themselves which is the smallest star and does it differ significantly from the largest planet? The current best answers are van Biesbroeck 8B (vB8B) and yes, according to speakers at a workshop* last autumn.

The difference between stars and planets is at least as much in how the objects form as in their present energy sources, and vB8B falls cleanly in the 'star' class, although contraction may be contributing as much to its luminosity as nuclear reactions do. I shall return shortly to the third big question addressed at the workshop: are there enough faint stars to affect estimates of the local mass density (one aspect of the widespread 'missing mass' problem)? To this, the democratically chosen answer was a resounding 'maybe', but the situation looks a good deal more promising for small stars than it did a year or two ago.

David C. Black (NASA/Ames) drew a sharp line between small stellar companions and planets¹. Stars form by fragmentation of a gas cloud without significant dissipation or chemical fractionation. Planets result when dissipation produces a disk around a single proto-star, within which bodies condense whose chemical composition and mass depend on distance from the centre of the disk. Alan P. Boss (Carnegie Institution, Washington) provided confirming evidence for this distinction. He has followed numerically the collapse and fragmentation of gas clouds of varying temperature and rotation rate until the fragments are so small that thermal support lets them contract slowly into single stars rather than breaking up further. No fragments smaller than 0.02–0.05 $M_{\odot}$ (20–50 Jupiter masses, M_J) ever formed, requiring a separate origin for planet-sized bodies. Most of the effects neglected in the calculation will tend to raise the limiting mass slightly. Boss and Hans Zinnecker (Royal Observatory, Edinburgh) both noted the main exception. Less dust lowers the limit by contributing less capacity, leaving open the possibility of still smaller stars within the metal-poor first generation, some of which may linger in the halo of our Galaxy.

None of the smallest fragments will ever burn hydrogen. This requires a mass $M \geq$

0.08 $M_{\odot}$ and provides the traditional definition of 'real' stars, presenting us with the problem of what to call the smaller objects. Jill Tarter (University of California at Berkeley) put forward a strong case for the continued use of the folk name 'brown dwarfs', on the grounds that objects whose emission spectra we cannot even roughly predict should properly be called by a name that is not, spectroscopically speaking, a colour.

Deciding whether there are enough brown dwarfs to contribute appreciable local mass density is much more difficult. The task has several pieces: (1) identifying faint-star candidates and making sure there are no systematic biases against detection; (2) acquisition of enough data (colours, distances) or accurate enough model atmospheres to determine absolute brightnesses or effective temperatures for the candidates; (3) converting these to masses; and (4) drawing plausible extrapolation curves without violating other known limits.

Identifying candidates is a struggle in itself, many projects yielding only upper limits. Bruce Campbell (Dominion Astrophysical Observatory, British Columbia) and Geoffrey Marcy (San Francisco State University) reported searches for variable stellar radial velocities that might reflect the gravitational influence of invisible companions. Both searches rediscovered known binaries, but the former ruled out companions bigger than 1–3 M_J with orbit periods less than 5 years for all half-dozen stars examined, and the latter excluded companions above 7–50 M_J with periods ≤ 1 year for all but one of 65 stars. The period limits are too short to tell us much about massive planets, but a possible class of brown-dwarf companions must be rare. Marcy's new find orbits Gliese 623 (CC 986) with a 2.1-year period and is probably 0.05–0.08 $M_{\odot}$. Curiously, Gliese 623 was already known, from astrometric and infrared studies, also to have a slightly more massive companion in a 3.7-year, non-coplanar orbit. An investigation of the stable lifetime of this system would be interesting.

Other non-detections include single brown dwarfs and companions to nearby white dwarfs in the Infrared Astronomy Satellite (IRAS) database reported by Frank J. Low (University of Arizona) and Harry L. Shipman (University of Delaware) and companions to nearby main sequence stars and white dwarfs in images recorded with the Infrared Telescope Facility (IRTF) by M.F. Skrutskie (Cornell University) and C. Krishna Kumar

* 1985 George Mason workshop on Brown Dwarfs, held at the George Mason University, Fairfax, Virginia 22030, USA from 14–15 October 1985.

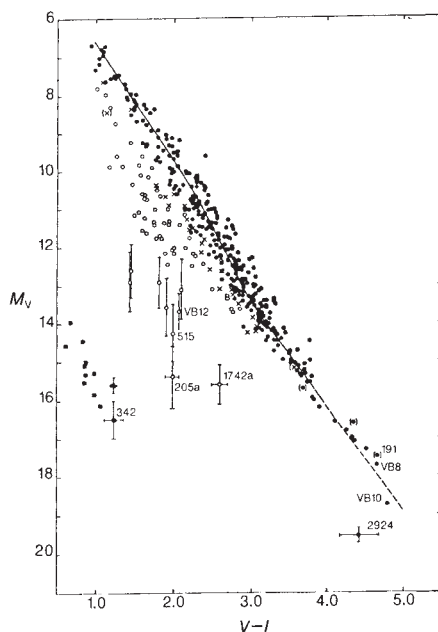
(Howard University). Because faint hydrogen burners, brown dwarfs and Jupiters all have radii near a tenth that of the Sun, these infrared searches essentially constrain brightness temperatures of hypothetical companions to be less than 2,000–2,500 K at the wavelength of observation (2–3 μm at the IRTF; 12 μm for IRAS). The uncertainties in turning these into luminosity and mass limits make interpretation difficult, except that the IRAS database should contain only about one identifiable brown dwarf even if such stars make up all the local missing mass and radiate like black bodies. Finally, Patricia C. Boeshaar (Rider College) discussed the status of a continuing search for faint dwarfs and subdwarfs based on multicolour (J,R,I) charge-coupled device images. So far, seven objects have turned up that fit on the bottom end of the subdwarf (low metal abundance) hydrogen-burning sequence, but no very red candidates have appeared.

Still, faint stars do exist. Lists shown by Ronald Probst (Kitt Peak National Observatory), James Liebert (University of Arizona) and Conrad Dahn (US Naval Observatory) included at least 16 with visual magnitude $M_v = 16$ –20, found, in most cases, by astrometry or in proper motion surveys and confirmed by measured distances and infrared fluxes (see figure). Most have the large velocities characteristic of old stars ($> 10^9$ years), and so must be hydrogen burners. A few seem to be young and could be contracting, cooling brown dwarfs.

Going still fainter, Robert S. Harrington (US Naval Observatory) has found eight astrometric binaries whose orbit solutions permit two kinds of companions, the comparable and the negligible. The former will have masses $\geq 0.08 M_\odot$ like their primaries and should be detectable infrared sources; the latter range downwards from 0.05 to 0.01 $M_\odot$ and should be very faint. Five of these have been examined by Donald W. McCarthy Jr (University of Arizona) as part of a continuing infrared speckle program that has already confirmed other astrometric companions as among the faintest stars. Four could not be seen, supporting the low-mass, brown-dwarf alternative. The detected fifth is vB8B, whose companion², with a 1.1–2.2 μm colour temperature of 1,360 K and luminosity $L = 3 \times 10^{-5} L_\odot$, is the coolest, faintest star known. It is part of a small stellar group whose dynamical age exceeds 2×10^9 years and so is almost certainly a hydrogen burner.

The greatest surprise among the faint-star searches came from M.R.S. Hawkins (UK Schmidt and Telescope Unit). His repeat of an earlier, rather discouraging search^{3,4} found 13 stars, rather than only 2, in the reddest bins, apparently because it went a bit deeper in the V colour band. Both his raw data and their transforma-

tion to absolute visual magnitude show a luminosity function that flattens off, begins to drop, but then rises again below $M_v = +16$. The 13 stars concerned have quite a low velocity dispersion, and Hawkins considers it possible that a new, young population of brown dwarfs is beginning to show up. The stars have been placed on a charge-coupled device parallax program for confirmation of their distances. Boeshaar expressed the opinion that these new counts are strictly inconsistent with her results, but given that the two projects



Hertzprung-Russell (colour-magnitude) ($V-I$) diagram of the nearby faint stars (C. Dahn and J. Liebert, personal communication). A clear, although sparsely populated, main sequence extends several magnitudes below $M_v = +16$. Some well-known stars are labelled, including vB8B. •, Disk dwarfs; x, mild subdwarfs; o, subdwarfs; ■, degenerates.

used slightly different colour bands and that neither presented completely uninterpreted numbers of images, many participants suspected that the two could be reconciled in the direction of a luminosity function that first drops and then rises again.

The increased number of faint candidates combines with several theoretical developments to raise once again the possibility that brown dwarfs may comprise all or most of the local (disk and/or halo) dark matter. First, the transformation from a brightness temperature to a luminosity is not entirely straightforward. Below 2,500 K, molecules become important and dust grains form, gradually raising the opacity, but also making it a very non-uniform function of wavelength. For instance, the flux in the frequently used wavelength band R is seen to be as much as 2.9 mag less than that of a black body of the same total luminosity among known faint stars, according to Probst. Thus some searches may still not be looking

quite faint enough, and we may be assigning wrong brightnesses to things we do find.

Second, clouds of grains will eventually merge to form a cloud deck, with photons reaching us only from the cool zones above it. This can happen rapidly enough, according to David J. Stevenson (California Institute of Technology) to cause a discontinuity in the evolution of brown dwarfs, in which they drop quickly from $\geq 2,000$ to about 1,300 K and then linger there for half a billion years or more while the trapped energy gradually leaks out. The corresponding luminosity is about that of vB8B ($3 \times 10^{-5} L_\odot$) which may, therefore, be in this plateau phase. The absence of slightly warmer objects in assorted searches would not be surprising, given this discontinuous evolution.

Finally, the simple power law commonly used for transformation from luminosity to mass turns out to underestimate severely the number of low-mass objects needed to account for a given luminosity distribution, $N(L)$. Francesca D'Antona (Osservatorio Astronomico di Roma) showed a new set of low-mass models⁵ that incorporate molecular and dust opacity as well as exact equations of state for partially degenerate material that burns hydrogen out of equilibrium. The main surprise is a great stretching of the main sequence down to low luminosities. This results from a class of transition objects of 0.07–0.08 $M_\odot$ that never quite burn hydrogen vigorously enough to stop contracting, but do have their contraction timescales stretched out to billions of years by the nuclear contribution. Thus, a very narrow mass range provides most of the stellar population over a wide luminosity range, 10^{-3} – $10^{-5} L_\odot$. The star vB8B falls in this range and is arguably such a transition object, of considerable age, and with a mass ($\sim 0.075 M_\odot$) near the upper end of the range permitted by the astrometric orbit.

Given the new models, the small number of known, faint stars implies a distribution of stellar masses, $N(M)$, that is at least flat or, more likely, still rising towards the brown-dwarf range. The high mass-to-light ratio of clusters of galaxies may still require massive theoryons, but the local problem begins to look much simpler, with ordinary faint stars, brown dwarfs and large planets as a possible solution. □

1. Black, D.C. & Levy, E.H. *Science* (submitted).
2. McCarthy, D.W., Probst, R. & Low, F. *Astrophys. J.* **290**, L9 (1985).
3. Reid, N. & Gilmore, G. *Mon. Not. R. astr. Soc.* **196**, 15 (1981).
4. Reid, N. & Gilmore, G. *Mon. Not. R. astr. Soc.* **201**, 73 (1982).
5. D'Antona, F. & Mazzitelli, I. *Astrophys. J.* **296**, 502 (1985).

Virginia Trimble is visiting Professor in the Astronomy Program, University of Maryland, College Park, Maryland 20742 and is Professor of Physics at the University of California, Irvine, California 92717, USA.

Major groove or minor groove?

SIR—A recent correspondent¹ re-emphasized that the major grooves of nucleic acid double helices provide more discriminating binding sites than do the minor grooves for proteins that recognize specific nucleotide sequences. Because of his mistaken impression that the major groove is accessible only in B-type double helices he went on to argue that proteins which have A-type DNA-RNA hybrids as binding partners would interact only with the less discriminating minor grooves.

It is true that A-DNA, the longest known² A-type structure, has a major groove that is clenched shut. This most compact A form also has the smallest axial component of nucleotide length ($h = 0.26$ nm). But the A family of double helices is quite polymorphic³ and less condensed allomorphs are observed usually under more hydrated conditions. They have nucleotide lengths nearly as large as in B-DNA ($h = 0.34$ nm) and major grooves that are wide — big enough to accommodate a third polynucleotide chain in some cases⁴.

Despite their substantial morphological differences, all the A allomorphs have the same types of rotations at each nucleotide bond. Little energy therefore can be needed to widen the major groove of any A-type helix to B-like dimensions. In these circumstances it is wrong to presume that the proteins involved in transcription, reverse transcription or priming of DNA synthesis cannot be involved with recognition sites in the major grooves of their substrates because of a structural constraint inherent in A-type double helices.

STRUTHER ARNOTT

Department of Biological Sciences,
Purdue University,
West Lafayette,
Indiana 47907, USA

1. Wang, J.-H. *Nature* **319**, 183–184 (1986).
2. Franklin, R.E. & Gosling, R.G. *Acta crystallogr.* **6**, 673–677 (1953).
3. Arnott, S. *et al.* *Cold Spring Harb. Symp. quant. Biol.* **57**, 53–65 (1983).
4. Arnott, S. & Bond, P.J. *Nature new Biol.* **244**, 99–101 (1973).

X chromosomes and dosage compensation

SIR—The suggestion that the primary function of X-chromosome inactivation in mammals might be one of sex-determination rather than dosage compensation made by Chandra¹ is a challenging one. Chandra questions whether dosage compensation is necessary, in view of Kacser and Burns' demonstration² that metabolic fluxes are insensitive to changes in enzyme activity of a factor of two. However, notwithstanding Kacser and Burns' work, the fact remains that all autosomal monosomies in the mouse (and also in humans) are

lethal at an early embryonic stage^{3,4}, and autosomal trisomies too are either prenatal lethals or result in malformation or retardation of young born alive³. Thus, it appears that for the X chromosome, with females having two copies and males only one, dosage compensation is indeed vital.

The original suggestion that X-chromosome inactivation is a form of dosage compensation therefore remains entirely reasonable. This seems particularly so in view of recent discoveries concerning the X–Y pairing segment. It was predicted long ago that the human X might have a pairing segment which was non-inactivated⁵. The grounds for such a prediction were obviously that if the function of X-inactivation was dosage compensation then a pairing segment carrying homologous genes would not require such a function. It is now known that the genes *Xg*, *Srs* and *MIC 2X* on the human X which lie in (*MIC 2X*)⁶ or near (*Srs*) the pairing segment do escape inactivation^{7,8}. Similarly, in the mouse, according to Keitges *et al.*⁹ the *Srs* locus is present on both X and Y, and is not inactivated.

However, although the evidence that the function of X-inactivation is dosage compensation seems compelling, Chandra's suggestion that it originally arose as a means of sex determination may still be valid, and provide a stimulating new view of the possible evolution of X inactivation. In lower vertebrates, as pointed out by Ohno¹⁰, the sex chromosomes are not heteromorphic. It is possible that in an ancestor of the mammals with homomorphic sex chromosomes, sex determination was achieved by inactivation of relevant genes in one sex. This inactivation might then have provided a basis from which dosage compensation might evolve, as the chromosomes became heteromorphic in the course of evolution of the mammals. The mechanism of X-chromosome inactivation in its present form appears to involve travel of some signal along the chromosome from an inactivation centre^{11,12}. If a mechanism of this type were present from an early evolutionary stage then the means would be available for inactivation to increase in extent so as to involve not only sex determining genes but also all loci in respect of which the X and Y had become heteromorphic.

In this connection the *Srs* locus is interesting. In the mouse, the locus is said to be present on both X and Y, and not inactivated⁹. In man, it is present on the X only, and although it is not inactivated the ratio of activities in XX and XY is not 2:1 but 1.6:1⁷, as a result of reduced activity of the allele on the inactive X. Could this be a transitional state? If so, there may be a third type of species with respect to *Srs* still to be discovered, so that there are, (1) mouse, X and Y homologous, no in-

activation, (2) man, X and Y non-homologous, partial inactivation, (3) unknown species, X and Y non-homologous, total inactivation. Another point to be expected from such a system of evolution of X-inactivation would be that originally the sex determining loci on the X would lie near the inactivation centre. Whether they would still do so today, in view of the rearrangements of the X that have occurred during evolution¹³ is uncertain, but this point may be worth bearing in mind when considering candidate loci.

MARY F. LYON

MRC Radiobiology Unit,
Chilton, Didcot,
Oxon OX11 0RD, UK

1. Chandra, H.S. *Nature* **319**, 18 (1986); *Proc. natn. Acad. Sci. U.S.A.* **82**, 6947–6949 (1985).
2. Kacser, H. & Burns, J.A. *Genetics* **97**, 639–666 (1981).
3. Epstein, C.J. *Trends Genet.* **1**, 129–134 (1985).
4. Magnuson, T. *et al.* *J. exp. Zool.* **236**, 353–360 (1985).
5. Lyon, M.F. *Am. J. hum. Genet.* **14**, 135–148 (1962).
6. Buckle, V., Mondello, C., Darling, S., Craig, I.W. & Goodfellow, P.N. *Nature* **317**, 739–741 (1985).
7. Shapiro, L.J. & Mohandas, T. in *Cytogenetics of the Mammalian X chromosome Part A* (ed. Sandberg, A.A.) 299–314 (Liss, New York, 1983).
8. Goodfellow, P.N., Pym, B., Mohandas, T. & Shapiro, L.J. *Am. J. hum. Genet.* **36**, 777–782 (1984).
9. Keitges, E., Rivest, M., Siniscalco, M. & Gartler, S.M. *Nature* **315**, 226–227 (1985).
10. Ohno, S. *Sex Chromosomes and Sex-linked Genes* (Springer, Berlin, 1967).
11. Lyon, M.F. in *Congenital Malformations* 67–68 (International Medical Congress, New York, 1963).
12. Rastan, S.J. *Embryol. exp. Morph.* **78**, 1–22 (1983).
13. Buckle, V.J. *et al.* *Cytogenet. Cell Genet.* **40**, 594–595 (1985).

Scotophobin resurrected as a neuropeptide

SIR—With the important role of neuropeptides in brain function and behaviour now established^{1–3}, it is interesting to review a piece of earlier work that never gained full credibility. Ungar and co-workers^{4,5} claimed to have isolated, sequenced, and synthesized a specific polypeptide, scotophobin. They further claimed scotophobin to be a memory molecule. Scotophobin was obtained from the brains of animals trained to avoid the dark (passive avoidance training) and was injected into untrained animals⁶. The naive recipients were reported to have changed their behaviour so as to avoid the dark. Ungar *et al.*⁵ concluded that acquired information at a molecular level had been transferred.

From the outset, the accuracy of Ungar and co-workers⁵ structure for scotophobin was questioned⁷, and more recent views include the idea that the artificial peptide might influence behaviour, but that no scotophobin may actually be synthesized in rat brain⁷. A review of the memory-transfer experiments more generally was made by Irwin⁸, who wrote that "From the beginning, the transfer paradigm became mired in an endless debate over the specificity of the behaviour allegedly transferred"⁸. Irwin went on to observe that "In the long run it was not

Table 1 Amino acid sequence comparisons

Scotophobin ⁵	(NH ₂)	YGGQQASKGQQNDS
Substance P precursor ¹¹	(88)	<u>Y</u> GHGQL <u>SH</u> KRHKTDS
Enkephalins ⁹	(NH ₂)	<u>Y</u> GGF ^M _L
β-Lipotropin ¹²	(61)	<u>Y</u> GGFMT <u>SE</u> KSQTPLV
MSEL-neurophysin ¹³	(63)	<u>S</u> GGRC <u>AA</u> FGVCCNDE

Numbers in parentheses give number of residues along polypeptide chain from amino terminal end. "NH₂" indicates amino terminal end. Underlined amino acids indicate sequence similarity to scotophobin.

confusion over the biochemical identity of the transfer factors . . . which most damaged the credibility of the transfer paradigm, but simply the unreliability of the phenomenon"⁸.

It is interesting to consider what these experiments might have, and might still contribute to the field of neuroscience. Ungar *et al.*^{4,5} appear to have presented early evidence of a peptidergic influence on behaviour, several years before the enkephalins were described⁹, but scotophobin is not mentioned in reviews of neuropeptides and is often omitted from descriptions of the history of development of this young field (for examples, see refs 1, 10). It does not appear that the identification and characterization of scotophobin contributed to the surge of interest in neuropeptides. Instead, one might suggest that the failure of the scientific community to accept the paradigm of molecular transfer of memory resulted in the scotophobin molecule being largely ignored.

With the now widely accepted view that certain neuropeptides play important roles as neurotransmitters and as modulators of brain function and behaviour^{1,3}, it might be worthwhile to consider scotophobin's possible role as a neuropeptide, more generally influencing behaviour as a neurotransmitter or neuromodulator. To that end, it is interesting to compare the amino acid sequence of scotophobin⁵ with that of other brain peptides and proteins (Table 1). The pentadecapeptide scotophobin has an amino-terminal peptide triplet that is identical to the enkephalins. Scotophobin also has sequence similarities to the precursor molecule for substance P. There appears to be an evolutionary relationship between substance P precursor and scotophobin, as judged by the stability of the sequence similarity under simulated evolutionary change (Protein Identification Resource program, $P < 0.01$).

Thus, scotophobin appears to be relat-

ed structurally to neuropeptides and their precursors, including polypeptides that play a role in the processing of information concerning pain in the nervous system. In that regard it is interesting to note that the animals in which scotophobin was said to have been induced underwent an extended and stressful training, including footshock. Induction of pain- or stress-related factors might be expected with such treatment.

Despite the uncertainty about the accuracy of scotophobin's structure, it might be worth reconsidering whether scotophobin or a related neuropeptide plays, or can play, a role in the modulation of behaviour in animals.

I thank the Protein Identification Resource (National Biomedical Research Foundation, Washington, DC) and Syed Pervaiz for his aid in accessing and using that resource.

DAVID WILSON

*Departments of Biology and
Physiology & Biophysics,
University of Miami,
Coral Gables, Florida 33124, USA*

1. Gainer, H. & Brownstein, M.J. in *Basic Neurochemistry*, 3rd ed., (ed Siegel, G., Roberts, W., Agranoff, B. & Katzman, R.) 269-296 (Little Brown, Boston, 1981).
2. Dismukes, R.K. *Behav. Brain Sci.* **2**, 409-448 (1979).
3. Zager, E.L. & Black, P.M. *Neurosurgery* **17**, 355-369 (1985).
4. Unger, G., Galvan, L. & Clark, R.H. *Nature* **217**, 1259-1261 (1968).
5. Ungar, G., Desiderio, G.M. & Parr, W. *Nature* **238**, 198-202, (1972).
6. Stewart, W.W. *Nature* **238**, 202-209 (1972).
7. Rainbow, T.C. *Neurochem Res.* **4**, 297-312 (1979).
8. Irwin, L.N. *Perspect. Biol. Med.* **21**, 476-491 (1978).
9. Hughes, J., Smith, T.W. & Kosterlitz, H.W. *Nature* **258**, 577-579 (1975).
10. Cooper, J.R., Bloom, P.E. & Roth, R.H. *The Biochemical Basis of Neuropharmacology*, 4th Edn (Oxford University Press, New York, 1981).
11. Nawa, H., Hirose, T., Takashima, H., Inayama, S. & Nakanishi, S. *Nature* **306**, 32-36 (1983).
12. Li, C.H., Barnafi, L., Chretien, M. & Chung, O. *Nature* **208**, 1093-1094 (1965).
13. Chauvet, M.T., Hurlbut, D., Chauvet, J. & Archer, R. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2839-2843 (1983).

Origin of anti-parallelism in β-keratin

SIR—Steinert and Steven¹, in a recent article on intermediate filaments, accepted as likely that the neighbouring pairs of two-chain coiled-coil molecules in the intermediate filaments of α-keratins are anti-parallel, and hence the filaments are non-polar structures. The primary evidence for an anti-parallel organization existing for the α-helices comes from the knowledge that the highly ordered β-crystalline state is derived from α-keratin. The β-chains have been shown to be primarily anti-parallel² and hence the α-helices, from which the β-chains were derived, presumably by extension, are also anti-parallel. It should, however, be noted that Menefee³ obtained in experiments on dipole disordering in fibrous keratins, a very definite net dipole vector for porcupine quill specimens which, on the basis of some elementary assumptions, he concluded gave a value of 38% by volume for

unidirectional oriented α-helical protein. In work on β-keratin derived by heating under pressure α-keratin fibres^{4,5} in water at 120–130°C, it has been shown that the α-keratin is converted to the β-state with little overall change to length of the keratin fibre.

The conclusion from these data^{4,6} was that the β-keratin structure resulted from the conversion of each α-helix into a β-chain folded into a single hairpin. Under these circumstances, the length change from the α-helix to the β-hairpin is about 10% compared with over 120% for the conversion of an α-helix to an extended β-chain. In the hairpin structure, the molecular chains on both sides of the hairpin are automatically anti-parallel.

In the case of β-keratin formation by the extension of an α-keratin fibre, there is little doubt that the β-keratin is formed in the extended state, for fibres extended in water at room temperature. However, to obtain their highly crystalline β-keratin specimens, Fraser *et al.*⁷ not only extended their α-keratin structure but also steamed the structure. This treatment, at elevated temperatures, conveys considerable mobility to the molecules. This mobility results not only from temperature activation, but also by the removal of the restriction on the molecular chains of the covalent interchain bonding of the cystine group via the process of sulphhydryl disulphide interchange⁸. This increase of mobility produced in steaming of the keratin structure could result in the folding of the elongated β-chains into β-hairpins.

The folding would be favoured by the lower energy state of the anti-parallel β-keratin as against the parallel β-structure². Similar folding has been observed in many long chain polymeric structures at a temperature just below the melting point of the polymer⁹. The possibility of anti-parallel β-chains being formed from parallel α-helices, certainly exists, with considerable physical evidence in its favour. Rather than suggesting¹ for α-keratin fibres that the 'intermediate filaments don't know whether they are coming or going', it may be more appropriate to suggest that we are all in this rather parlous state.

M. FEUGHELMAN

*University of New South Wales,
PO Box No. 1,
Kensington, New South Wales,
Australia 2033*

1. Steinert, P.M. & Steven, A.C. *Nature* **316**, 767 (1985).
2. Fraser, R.D.B., MacRae, T.P., Parry, D.A.D. & Suzuki, E. *Polymer* **10**, 810-826 (1969).
3. Menefee, E. *Ann. N.Y. Acad. Sci.* **238**, 5367 (1974).
4. Feughelman, M. & Mitchell, T.E. *Biopolymers* **6**, 1515-1518 (1968).
5. Mitchell, T.W. & Feughelman, M. *Kolloidzeitschrift und Zeitschrift für Polymere* **229**, 124-131 (1969).
6. Feughelman, M. *Textile Res. J.* **40**, 1125-1126 (1970).
7. Fraser, R.D.B., MacRae, T.P. & Rogers, G.E. in *Keratins* (Kugelmas, I.N. ed.) 114 (Thomas, Springfield, Illinois, 1972).
8. Feughelman, M. *Proc. Third Int. Wool Text. Res. Conf., Paris 1965*, **2**, 245-257 (Institute Textile de France, Paris, 1965).
9. Lindenmeyer, P.H. *J. Polymer Sci.* **20C**, 145-158 (1967).

A place for teleology?

William H. Press

The Anthropic Cosmological Principle. By John D. Barrow and Frank J. Tipler. Clarendon: 1986. Pp.706. £25, \$29.95.

TELEOLOGY, we recall, is the doctrine that natural phenomena are guided not only by immediate causal forces, but also by certain pre-determined, distant goals. It is an understatement to say that teleology has been out of fashion in science in the past century, and probably not much of an exaggeration to say that the present scientific paradigm rejects the teleological hypothesis vehemently, categorically and usually with contempt. If the Darwinian revolution was not about the rejection of teleology, then what was it about?

Now there comes a book — no crackpot tract, but a scholarly, philosophically sophisticated and mathematically high-brow monograph — that says we've all made a big mistake; there really is a place for teleology and related concepts in today's science. At least (the authors ask), give us a chance to present arguments, drawn in the main from modern theoretical cosmology, which may convince the reader of an astounding claim: there is a grand design in the Universe that favours the development of intelligent life. This claim, in certain variations (as we shall see in a moment), is the "anthropic cosmological principle".

The immediate dilemma is whether or not to take such a book seriously at all. There are several good reasons to do so. First, it is a magnificently quixotic quest that the authors are undertaking, and they approach it with style, erudition and panache, though not without some sly reliance on debaters' tricks. Second, as the rest of us putter around in the basement of currently accepted scientific beliefs, it is not a bad thing to be prepared to defend those beliefs — and this book surely offers a dangerous (but in the end, I think, rebuttable) challenge to them.

Third, although Barrow and Tipler try valiantly to cloak their thesis in the garb of heavy philosophy and impenetrable (and sometimes misused) mathematics, there lurks within the book a delightful, wide-ranging and provocative exploration of their own hodge-podge of broad intellectual interests. These centre on modern relativistic cosmology (in which the authors do hold established credentials) and range from there with enthusiasm, and greater or lesser competence, into biochemistry

and the origin of life, into space travel and extraterrestrial intelligence, into many-world interpretations of quantum mechanics, into critical analysis of the work of Jesuit mystic Teilhard de Chardin and the Humean rejection of design arguments for the existence of God, and much, much more. It would be quite a challenge to find another book whose "B" index entries include Boshongo creation myths, the auto-



catalytic role of beryllium-8 in the helium-burning reaction and Bianchi type VII universes.

So what business have a couple of mathematical cosmologists poking around in places such as this? Barrow and Tipler have found themselves a kind of intellectual magic carpet, which transports them effortlessly through parochial disciplinary barriers into any field they want to get into. The anthropic principle is that magic carpet. Under its present name, it has been talked about by (mostly) cosmologists for something more than a decade; Barrow and Tipler trace it, and its forerunners and precursors, back through several centuries.

A weak, and not very controversial, form of the anthropic principle is the notion that observational selection biases should be accounted for in cosmology, just as they are accounted for in less grandiose fields of observational or natural science. As the authors put this:

The observed values of all physical and cosmological quantities are not equally probable but they take on values restricted by the require-

ment that there exist sites where carbon-based life can evolve and by the requirement that the Universe be old enough for it to have already done so.

In other words, we know that *we* are here (Bishop Berkeley, Descartes: these authors have you all figured out), so we might as well rule out physical theories and cosmological models that are incompatible with that "observation". Taken this way, the weak anthropic principle (WAP) is little more than an application of Bayes theorem of probabilities, letting a little bit of *a posteriori* knowledge (the existence of life on Earth) intrude into an assessment of the likelihood of *a priori* theories.

There is only the faintest whiff of teleological dogma in the WAP. By accepting it, one gains intellectual dominion over a whole set of territories which would otherwise lie outside the scope of science. "Why" do the physical constants have the values that they do? "Why" is the Earth the size and temperature that it is? And so on. The meat of this book is the exploration of consequences of the WAP in cosmology, astrophysics, planetary physics and evolutionary biology. But the seductive trap that the authors are setting is already clear: as soon as we accept the "why" formulation of questions that the WAP allows us to address, we have entered a receptive state of mind for the strong anthropic principle (SAP): "The Universe must have those properties which allow life to develop within it at

some stage in its history". Or, more extreme yet, John Archibald Wheeler's participatory anthropic principle (PAP): "Observers are necessary to bring the Universe into being". Barrow and Tipler coyly avoid saying that they want us to believe in one of these stronger statements, but one cannot read their book without feeling that they do, that they are earnest, wide-eyed, proselytizing believers themselves. One can almost see them handing out flowers in airports or train stations.

It is at this point that, not without reluctance, I have to bail out of their plane and parachute to safety. Barrow and Tipler have given us an engaging book, practically a universal education in both the history of modern science and the history of the Universe. (In a short review one cannot do justice to the diversity of their topics and the fascinating historical titbits that they have unearthed and recorded.) This book will be much quoted, much debated and much praised. It deserves a place on the shelf of any serious scholar of science, and it will be found on a

good many coffee tables as well. Nevertheless, in my opinion, there is some fundamental intellectual dishonesty here, some snake oil being peddled. The authors are trying desperately to be taken seriously, and they do not always play fair with their readers.

While there is a lot of good scientific exposition in the book, there is too a distressing amount of what seems to be mathematical flim-flam, that is, quotation of precise results in a manner designed to mislead less-mathematical readers and cause them to jump to the authors' desired (usually non-mathematical) conclusion. For example, after some rather woolly discussion of Penrose conformal diagrams, Barrow and Tipler conclude "A Penrose diagram allows us to define rigorously 'an achieved infinity', a concept whose logical consistency philosophers have been doubtful about for thousands of years". This is a silly assertion, but it is put forth with the utmost gravity, in such a way that many readers will be taken in. And it is only one of many such cases.

Nor do the authors always play fair with other workers in their field. They have a tendency to bury footnote references to the work of their colleagues in and among a plethora of erudite footnotes to nineteenth- and twentieth-century precursors. This becomes particularly disturbing when, as is sometimes the case, the exposition mirrors other authors' papers practically sentence by sentence. The casual reader will come away mistakenly crediting Barrow and Tipler for ideas that they have, in fact, only nicely summarized. The serious scholar should take their bibliographical research and attributions (an immense amount of work on their part) as indicative, not definitive.

The book's flaws, I think, can be traced back to a single cause: the authors badly want to be the founding doctrinal theorists of a "new" resurgence of teleological belief in science. They bring to their cause an impressively broad knowledge of scientific exotica, but the factual content of the book is only their means to an end that (in my opinion) is threatening to the modern scientific enterprise. That end is nothing less than the fusion of matters of science with matters of individual faith and belief. It has taken us a long time to separate these matters, each to its own legitimate arena in human affairs. We should not lightly allow them to become once again jumbled, least of all by a book that now and then cuts corners in getting to its own pre-determined, distant goals. This is a fascinating and entertaining book, one to read and think about. But it is also one whose extra-scientific agenda most of us will, ultimately, wish to reject. □

William H. Press is Professor of Astronomy and Physics at the Center for Astrophysics, Harvard University, 60 Garden Street, Cambridge, Massachusetts 02138, USA.

What do intermediate filaments do?

Keith Burridge

Intermediate Filaments: A Review. By Peter Traub. Springer-Verlag: 1985. Pp. 266. DM 168.

FOR SEVERAL years people interested in intermediate filaments have looked over their shoulders with envy at those studying microtubules and microfilaments. While there is ample evidence for the importance of microtubules and microfilaments in generating cellular movements, the function of intermediate filaments remains elusive. Because no obvious motility function has been apparent, some sort of skeletal or structural role has been generally, if begrudgingly, accepted, but this concept has not been embraced with the kind of enthusiasm that gives a field new momentum.

Anyone examining the function of intermediate filaments has some puzzling experimental observations to consider. Some cells exist, apparently quite happily, without them. Similarly, the intermediate filaments within a cell can be collapsed by microinjection of antibodies without detectable effects on the cell's behaviour. From these observations one might conclude the intermediate filaments are not very important, or, alternatively, that they have a subtle function which has been missed because the assays have been inappropriate. This latter argument is the one adopted by Peter Traub. He proposes that the idea that intermediate filaments have a structural role is essentially the consequence of depending on structural techniques, such as immunofluorescence and electron microscopy, to study them. Traub uses the final chapter of his book to present his own, novel ideas about possible functions; rather than being skeletal elements within cells, he suggests, intermediate filaments are involved in signal transduction from the plasma membrane or cytoplasm to the nucleus.

Traub first became entangled in this field accidentally. Working as a virologist, he became intrigued with a protein that was binding to some viral RNA. Only later did he realize that this protein had been discovered previously and was vimentin, the subunit of one of the main types of intermediate filament. At this point others might have dropped the subject, attributing the binding to irrelevant coincidence, but Traub and his colleagues have pursued the matter. They have shown under physiological conditions that the subunit proteins of several classes of intermediate filament have a specificity for binding to DNA rather than to RNA. From the work of others, it has been

known for some time that intermediate filaments are very susceptible to proteolytic cleavage by a calcium-activated protease in the cytoplasm, and again Traub has demonstrated that the fragments generated from the different classes of filament by this protease retain the selective DNA-binding activity.

These biochemical observations form the basis of Traub's theory. He suggests that calcium elevated in the cytoplasm in response to external signals activates the protease. This, in turn, cleaves the intermediate filament subunits, which can no longer assemble, and which are transported back to the nucleus. Here they interact with DNA or chromatin, probably in conjunction with other proteins, in a regulatory manner. When I first read about this idea I was sceptical. I still am, but Traub assembles a considerable body of evidence that is consistent with the proposal. Certainly, an advantage of his hypothesis is that it suggests many experiments by which it could be tested. On the other hand, the idea that intermediate filaments have primarily a skeletal function is a difficult concept to establish or to destroy.

Only the last chapter of the book is dedicated to this unorthodox but refreshing concept. Regardless of whether or not one accepts it, the book as a whole provides an excellent review of intermediate filaments. Almost every aspect of the subject is covered, with references up until about 1984, ranging from heterogeneity, structure and assembly, to the interaction of intermediate filaments with other elements, secondary modification and response to normal and pathological stimuli. Given how vast and labyrinthine the field has become, this is quite an accomplishment. The only area in which Traub has not attempted to be comprehensive is the increasing clinical use of intermediate filaments for tumour typing.

The book should be useful both for specialists and for the many cell biologists who, like myself, have a peripheral interest in intermediate filaments. It is well organized, well indexed and has a bibliography of some 60 pages. I anticipate that it will be the standard work for many years. If it also stimulates some novel experiments that help to unravel the knot of intermediate filament function, then Traub should be well pleased with having provided a double service to the field. □

Keith Burridge is in the Laboratories for Cell Biology, Department of Anatomy, University of North Carolina at Chapel Hill, 111 Swing Building 217H, Chapel Hill, North Carolina 27514, USA.

• Two recent books on the cytoskeleton are *Molecular Biology of the Cytoskeleton* (G.G. Borisy et al., eds), published by Cold Spring Harbor Laboratory, and *The Cytoskeleton: An Introductory Survey* (by M. Schliwa), published by Springer-Verlag.

Temper out of time

J.S. Jones

The Evolutionary Process: A Critical Review of Evolutionary Theory. By Verne Grant. Columbia University Press: 1985. Pp. 499. \$40.

ONE of the pleasures of being an evolutionist is that one's colleagues are such a cantankerous lot. They have, of course, a great deal to be cantankerous about. Each one of Darwin's precepts — that there exists extensive genetic diversity which reflects variation in fitness, that most polymorphism is subject to natural selection and that the gradual accumulation of favoured variants leads to the origin of species — has been disputed since they were first stated, and molecular biology has given new heart to those whose chief pleasure is in attacking the Darwinian edifice.

This book treats evolutionary theory from the viewpoint of an author who has grown up with the subject. Grant comes from that generation of biologists, now almost extinct, which sees evolution in the context of a deep understanding of the world of living animals and plants rather than as experts in a newly invented (and somewhat nebulous) science of "evolutionary biology". He discusses many of the classics of the field: gene flow, population structure and patterns of speciation in *Drosophila*, industrial melanism, genetic load, adaptive radiation and human evolution. Many of his examples come from plants, with sequoias as exemplars of population structure, the origin of corn as an instance of the power of selection and the annual herb *Gilia* as an illustration of reproductive isolating mechanisms.

There is, too, space for the grinding of a few private axes. Kin selection has, apparently, only slight explanatory power and a great ability to cause confusion of thinking. The same is true of speciation in continuous populations. However, many of the book's themes are acceptable to the most conventional evolutionist: the vast majority of new mutations are deleterious, natural selection always has the last word in determining the fate of a newly arisen variant and macroevolution can be explained as the summation of many microevolutionary events.

For all the strengths of a treatment of this sort, the book suffers from its concentration on convention. Traditional explanations will, no doubt, explain many of the patterns of morphological evolution which we see in nature. However, most of the recent excitement in evolutionary theory has arisen from the discoveries of molecular biology. It is now clear that there is a staggering amount of genetic diversity in DNA structure and that, at

this level, some of the patterns of genetic differentiation among closely related species are very difficult to explain in terms of mutation and selection. Grant's subtitle is *A Critical Review of Evolutionary Theory*, but it is not easy to be critical in a review which omits many of the ugly facts which now threaten Darwin's beautiful theory. Molecular evolution does get a mention, but its wider implications are scarcely discussed. As a result the book conveys little of the genuine bad temper which is such an attractive feature of modern evolutionary biology.

This, then, is a gentleman's book, which deals largely with the controversies of yesterday. It may well be that these will be the controversies of tomorrow; but what is needed is a fuller discussion of their relevance to the discoveries of today. □

J.S. Jones is a Lecturer in the Department of Genetics and Biometry, University College London, 4 Stephenson Way, London NW1 2HE, UK.

Views of the virus

P.J.G. Butler

Virus Structure and Assembly. Edited by Sherwood Casjens. Jones and Bartlett: 1985. Pp.295. \$45, £45.

VIRUSES are studied not only as interesting microorganisms in their own right, but also for what they can tell us about the host cells which they infect. Because many of them have small genomes, coding for few proteins (sometimes only three, including the coat protein for packaging new virions), they have to suborn many cellular functions for synthesis and even sometimes assembly of their components. Even in viruses with large genomes, such as bacteriophage T4 which codes for many enzymes, a fair proportion of the genes duplicate functions already available in the host cell and are inessential except in unusual hosts. Thus a virus-infected cell carries out many of its normal functions, but to a grossly exaggerated extent, and with a narrower range of products, which facilitates research into these processes.

There are many reviews of specific areas of virus structure and assembly aimed at specialist audiences, but no attempt has been made recently to cover the whole field in a single volume. A book such as the present one is therefore to be welcomed and it is all the more regrettable that it is in the area of all-round coverage that *Virus Structure and Assembly* is at its weakest; it is, perhaps, the difficulty of covering such a vast amount of information which has deterred others from attempting the task. The editor has gallantly contributed the two general chapters, "An Introduction to Virus Structure and Assembly" and "Nucleic Acid Packaging", himself. The second of these occupies a full quarter of the book, but despite its relative length I feel that it falls short of the mark. Inside Casjens's own areas of expertise the subject is well handled, but in dealing with less-familiar material he has resorted to a simple listing of published views with little attempt to assess their merits or significance.

The book contains several excellent chapters, which are both enjoyable and informative, but in the main these cover areas close to the editor's research interest (that is, bacteriophage). While I must declare an interest in the topic, it is inadequate to try to summarize tobacco mosaic virus, for which the most detailed information for any virus is available on both structure and assembly, in just four pages. In comparison, an entire chapter is devoted to the tailed bacteriophages, which also take large parts of two further chapters. While they demonstrate the power of genetic analysis, we have little understanding of their detailed interactions as there are no high-resolution structures known.

On the topics which are given adequate space this book does a good job, but it fails to meet the objective which the editor sets out in his preface; it would be wrong for anyone outside the field to be deceived by the title into thinking that it gives an overall picture of our knowledge of virus structure and assembly. It is, nonetheless, certainly worth reading. □

P.J.G. Butler is in the Medical Research Council's Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

Ambiguous shapes: these two figures are each open to different interpretations according to whether they are mentally rotated 45° clockwise or 45° anticlockwise (one of them, for instance, can be seen as a man in a chef's hat or, alternatively, as a dog).

The illustrations are reproduced from the new paperback edition of Mental Images and Their

Transformations, by R.N. Shepard and

L.A. Cooper, published by MIT Press price \$9.95, £9.95. For review and discussion of the issues raised by this line of research see Nature 301, 353 (1983).

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

An image of science and business

Peter Newmark

A Machine Called Indomitable. By Sonny Kleinfeld. *Times Books, New York:1985.* Pp.250. \$16.95.

RAYMOND Damadian, the main subject and inspiration for this book, could, it seems, have chosen to make a career as a violinist or a tennis player had he not chosen to study science and later medicine. By the end of Sonny Kleinfeld's rather shallow account, his subject has become a business man, grown rich by selling the nuclear magnetic resonance (NMR) scanner that he had pioneered. But this is much less a story of how the son of some poor Armenian immigrants made good in the great American tradition, than how a misfit in the conventional world of medical research was forced into business to support himself.

Damadian's problems with his fellow researchers seem to have been catalysed by his early association with Gilbert Ling, a physiologist who believes that the interior of cells comprises multiple polarized layers of water molecules which account for phenomena such as those otherwise interpreted as a sodium pump in the cell membrane. Falling under the spell of Ling — who remains a voice in the wilderness — Damadian first reasoned that the water structure of cancer cells would differ from that of normal cells, and then came across NMR as a means of detecting that difference: "One afternoon, peering at the pool... Damadian went off into his own mind. Gradually, swirling like smoke, visions emerged. Perhaps some sort of giant NMR machine might be just the device that could be used to scan his stomach (which was still acting up)..." (p.27). In the style of the reporter (for the *New York Times*) that he is, Kleinfeld recounts how the dream of 1970 became the prototype machine of 1977, and the business of today.

Damadian's first "mistake" was to outline in *Science* his vision of using NMR to detect tumours on the basis of their abnormal water structure. The paper seems to have been the subject of a good deal of derision, from which its author never really recovered in the eyes of many of his peers. The result was a continual struggle for money to support his work and a growing contempt and fear of his rivals, especially Paul Lauterbur at the University of New York at Stony Brook, whose paper on NMR imaging (for which he coined the unused term *zeugmatography*) was published in *Nature* in 1973. Damadian was incensed that the paper contained no reference to his own work; he was, however,

inspired by Lauterbur's achievement to switch his own efforts to imaging.

Perhaps Damadian's character is best illustrated by his financial struggles. When his first grant application was turned down he wrote directly to President Nixon: "The rejection of my grant by the National Institutes of Health is a colossal stupidity [which] reflects only the enduring ignorance and the complete absence of vision of the men who decide". Urged by NIH to reapply, he received a three-year grant with which he bought his first NMR machine. When that grant ran out and the "appallingly disgusting" peer review system failed him again, he raised money privately with the help of a brother-in-law

Reasoning that his fatness was the problem, he put pressure on his only slim associate to be scanned by the machine — by then christened *Indomitable*. Persuasion took two months, the scan itself four and three-quarter hours. But the result was an image. And a press conference. And more hostility.

Failing to get recognition, promotion or papers published, and once again without the money to proceed, it was then that Damadian decided to go into business. He also decided to switch from a superconducting magnet, with its problematic requirements for cooling by liquid helium, to a permanent magnet, which weighs 3,000 g in the latest machines. They now

Tumor Detection by Nuclear Magnetic Resonance

Abstract. *Spin echo nuclear magnetic resonance measurements may be used as a method for discriminating between malignant tumors and normal tissue. Measurements of spin-lattice (T_1) and spin-spin (T_2) magnetic relaxation times were made in six normal tissues in the rat (muscle, kidney, stomach, intestine, brain, and liver) and in two malignant solid tumors, Walker sarcoma and Novikoff hepatoma. Relaxation times for the two malignant tumors were distinctly outside the range of values for the normal tissues studied, an indication that the malignant tissues were characterized by an increase in the freedom of tissue water.*

Cause of controversy: the opening lines of Damadian's paper in *Science*, 19 March 1971.

through a Citizens' Campaign for New Approaches to Cancer. There followed a large National Cancer Institute grant (and a patent) but again no renewal. Later, in desperation, Damadian travelled to Plains, Georgia, to try and persuade the newly-elected President Carter of his case. Hugh Carter, a cousin, did his best on Damadian's behalf but no money was forthcoming. In the end it was a wealthy businessman who came to the rescue.

Despite financial problems, Damadian was making progress in the construction of a human NMR scanner. But the lack of money meant that much labour and ingenuity had to be spent on building everything, including the superconducting magnet, from scratch and often from junk. He was driven to a considerable extent by the fear that Lauterbur, or one of the British groups that had entered the field, would be the first to produce NMR images of the human body, in which case Damadian foresaw that he would lose any remaining chance of getting the credit for his invention. Lauterbur, however, seems not to have been concerned. A research associate at the time is quoted as saying "He didn't think Damadian was capable of producing an image. He thought he was some crazy doctor".

Whereas Lauterbur and others were busy imaging lemons, green peppers and even rats, Damadian was going for broke. Finding no volunteer among his colleagues to be the first human to be NMR scanned — there was some apprehension about the possible effects of a long exposure to a strong magnetic field — Damadian became his own guinea pig in May 1977. The result was not even a lemon.

sell for more than \$1.6 million and have to compete with similar machines from at least a dozen rival companies.

Kleinfeld's book has its attractions. It provides a plausible picture of a man who suffered the results of allowing his enthusiasm to get the better of him when the rules of the game demand a more measured approach. It is easy to see both why he became almost paranoid about his peers and why they took him less than seriously. It is, however, hard to be too enthusiastic about a writer who can say (p.8): "Two of the important substances that get reabsorbed [in the kidney] are sodium and potassium; they are the secret to biological electricity, which is the foundation of life. It allows us to move our arms and legs and to breathe". Also, even though the book does not pretend to be a balanced account of the development of NMR imaging, it would have benefited from a less Damadian-dominated version of the story. The author talked to Lauterbur and some of the British rivals, but one learns little of the former and next to nothing of the latter.

Ironically, NMR scanners are still of little use in detecting tumours, the point of departure for Damadian's quest. Undaunted, Damadian now foresees the development of *Indomitable* into a tumour-killing device. And, indomitably, he still believes that in cancer "the real event that's taking place is that the water is becoming more and more disorganized and less structured in the cell". For better or for worse there is probably more room for such characters in business than in science. Or even in music or tennis. □

Peter Newmark is Deputy Editor of *Nature*.

The management of sea-level rise

from Walter S. Newman and Rhodes W. Fairbridge

Man can regulate the hydrological cycle by diverting ocean-bound fresh water into reservoirs, subsurface storage and interior basins, and by diverting sea water into continental depressions. Between 1957 and 1982, human intervention stored as much as 0.75 mm yr^{-1} of sea level rise potential in reservoirs and irrigation projects.

TIDE gauge studies reveal that the trend of sea level in this century has been generally upwards along most of the world's coastlines. For most of this century, the mean rise has been more than 1 mm yr^{-1} (refs 1–4). Meier⁵ believes that the wasting of mid-latitude glaciers by 'greenhouse-effect' warming is in part responsible for the modern sea-level rise, whereas Gornitz *et al.*¹ find the 'steric effect' (expansion of near-surface water due to temperature increase) explains much of the remainder of the current rise.

In the past few years, several independent reports (for example, refs 6, 7) have presented rather alarming scenarios predicting a greenhouse-effect sea-level rise on the order of 0.6 to 3.5 m by the end of the next century⁸. Although the CO_2 models may be controversial, the tide-gauge evidence is qualitatively correct. Whatever might be the ultimate cause of the contemporary sea-level rise, the continuing marine inundation poses a major threat to the world's littoral population.

Options

The problem can be addressed in several ways: by evacuation and retreat from the contemporary littoral (including legislated withdrawal from coastal habitations and investments), defence of the shore by a variety of engineering schemes, as practised, for example, in the Netherlands, and/or manipulation of the hydrologic cycle. On a worldwide basis, we have no means of estimating the cost of either the evacuation or the defence options. For most littoral states, either of these options would presumably consume an appreciable portion of their gross national products.

We examine here a third option: anthropogenic management of the hydrologic cycle. A eustatic sea-level rise of 1.25 mm yr^{-1} equals an annual increase in ocean volume of nearly $500 \text{ km}^3 \text{ yr}^{-1}$. We assume that these figures are a conservative approximation of the observed trends. The crux of the problem of regulating sea level is then how to interrupt the hydrological cycle so that each year we can intercept and dispose of, or store, $\sim 500 \text{ km}^3$ of additional water volume apparently being added to the ocean mass.

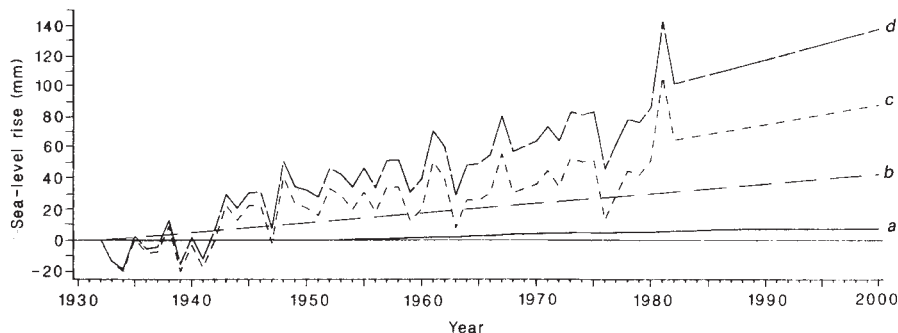


Fig. 1 Sea-level volume suppression since 1932 due to reservoir and irrigation project storage. The global sea-level plot (c) to 1982 is based on data by Gornitz and Lebedeff²²; thereafter an assumed volume increase of $500 \text{ km}^3 \text{ yr}^{-1}$ is used to extrapolate the sea-level rise to the year 2000. Trace a shows the growth of the known storage capacity of the world's 107 largest reservoirs⁹ (capacity $\geq 1 \text{ km}^3$); b is our estimate of water volume detention in small reservoirs and irrigation projects. Trace d adds the anthropogenically-stored water (a plus b) to trace c, to generate what would have been the trace of sea level unsuppressed by continental storage. Note that this trace does not include the effect, if any, of groundwater mining.

In fact, we already store an appreciable volume of water on the continents. The Soviet hydrologist M. I. L'vovich⁹ estimates that mankind now regulates 15% of the Earth's annual river discharge. Figure 1 (trace a) shows the rate of growth of the water storage capacities of the world's 107 largest reservoirs¹⁰. The total quantity already stored or planned for storage by the end of the century is $\sim 3,100 \text{ km}^3$. These data imply that since 1932, sea-level rise has been suppressed by reservoir storage.

Figure 1 also shows the 'noisy' rise of sea level since 1932 (trace c). This trace is derived from an updated and expanded database of tide-gauge measurements and late Holocene sea-level indicators assembled by Gornitz and Lebedeff²² and used by them to obtain a global average sea-level rise of $1.2 \pm 0.3 \text{ mm yr}^{-1}$. We used their data for the interval 1932–1982 and extrapolated the slope of their data (1.25 mm yr^{-1}) to the year 2000.

The water stored in large reservoirs would otherwise have been responsible for an appreciable annual increment in modern sea-level rise. For the time interval 1957–82, the mean rate of suppression has been about $125 \text{ km}^3 \text{ yr}^{-1}$ (equal to a potential rise in sea level of 0.25 mm yr^{-1}). This volume does not include the construction and closing of millions of smaller reservoirs which might double the rate of sea-

level-rise suppression to 0.5 mm yr^{-1} ($250 \text{ km}^3 \text{ yr}^{-1}$). Nor does this volume include irrigation water, some of which infiltrates into aquifers, although at least half of it is transpired and evaporated and most of the remainder runs off again towards the world ocean¹¹. We estimate that about $125 \text{ km}^3 \text{ yr}^{-1}$ of some $3,000 \text{ km}^3 \text{ yr}^{-1}$ of consumed irrigation water infiltrates to the subsurface, or remains as soil moisture and is more or less permanently stored in association with irrigation projects. We suggest that since 1932, all forms of reservoir storage plus irrigation projects have been storing $\sim 375 \text{ km}^3 \text{ yr}^{-1}$ of water, equivalent to a potential sea-level rise of $\sim 0.75 \text{ mm yr}^{-1}$. The estimated rate of small reservoir and irrigation storage since 1932 and extrapolated to the year 2000 is shown as trace b in Fig. 1.

In contrast to water currently being withdrawn from the hydrological cycle is water formerly in storage but now being recycled. The USSR Committee for the International Hydrologic Decade has estimated¹² a sea-level rise of 0.8 mm yr^{-1} due to decrease in groundwater storage, and Gornitz *et al.*¹¹ believe that land storage is probably not a major cause of sea-level change because new storage is being balanced by groundwater mining. Both of these estimates are based on extrapolations of water mined from the Great Plains

Ogallala aquifer that extends north from western Texas through to the Dakotas. However, the lateral transmissivity of the Ogallala is very low because of the variability of the sediments in this alluvial formation¹³. It is therefore probable that their estimates of withdrawal are excessive by an order of magnitude or more. For example, Nace¹³ wrote that "Pumpage from the Ogallala formation has reduced storage by nearly 110 km³, and annual withdrawals in recent years have been about 6.2 km³." Furthermore, Ambroggi¹⁴ insists that the annual rate of irrigation storage clearly exceeds groundwater mining in arid areas. The total quantity of mined groundwater is difficult to determine, but probably amounts to less than 100 km³ yr⁻¹. Groundwater mining, if real and appreciable, would require a subtraction from reservoir and irrigation storage water accumulation in Fig. 1.

Trace *d* in Fig. 1 shows the result of adding the storage capacities of large and small reservoirs and irrigation projects (traces *a* and *b*) to the observed and extrapolated sea-level rise since 1932 (trace *c*). The result is a sea-level curve as it might have evolved had it not been for the several classes of anthropogenic continental water storage. The difference between the observed and projected sea-level curves *c* and *d* represents our estimate of the suppression of the 'true' sea level trajectory by continental water storage.

In addition to freshwater storage, it is also feasible to channel sea water into some of the world's great interior drainage basins which lie well below present sea level. Such potential sinks include the Imperial Valley of California, the Qattara Depression of northwestern Egypt, the Dead Sea rift valley of Israel and Jordan, the Salina Gaulicho of Argentina and the Eritrea (Lake Assale) depression in Ethiopia. All of these schemes also possess appreciable hydroelectric power potential. Marine diversion engineering feasibility studies already exist for the Dead Sea Basin and the Qattara Depression^{15,16}. We estimate the storage potential of each of these depressions to be on the order of 1,000 km³ or more.

The site with the greatest potential for freshwater storage is the Aral-Caspian depression, which lies within the Eurasian landmass. The Caspian Sea surface is at an elevation of -28 m and falling. Raising the level of the Caspian Sea by only 10 m would store 4,420 km³ of water, which represents eight years of sea-level rise at 1.25 mm yr⁻¹. The Soviet Union is already diverting water from the north-flowing Vychegda and Pechora rivers in European Russia, but this will serve only to reduce somewhat the rate of fall of Caspian Sea level^{17,18}. The three northern Siberian rivers (Ob, Yenisey and Lena) discharge about 1,500 km³ yr⁻¹. Soviet engineers are

planning the diversion of these rivers towards the south for irrigation and hydroelectric power as well as the maintenance of the Caspian and Aral Sea levels¹⁹. Additional water will be retained in smaller upstream reservoirs as well as in soils and subsurface aquifers. The central Eurasian project is thus certainly a key element in sea-level management.

Still other major and minor sinks offer considerable water storage potential and other benefits. Examples include the Great Artesian Basin of Queensland and New South Wales in Australia, and the middle Niger and upper Zaire valleys (with spillways into the Chad Basin) in North Africa. The several-decades-old NAWPA (North America Water and Power Alliance) plan to divert surplus water from Canada into the United States²⁰ would store about 1,000 km³ of fresh water in the Rocky Mountain Trench, equal to ~2.0 yr of sea-level rise at 1.25 mm yr⁻¹. The plan also envisaged additional storage by the damming and partial revival of ancient proglacial lake basins and recharging of subsurface water aquifers. With water and sediment loading, isostatic downwarping of all these reservoirs would gradually increase their capacity by an estimated ten per cent, although some seismic risk is always involved with such continental crustal loading.

Consequences

The major benefit of sea-level regulation is the mitigation of the damage inevitably caused by continuing littoral zone inundation. A cost-benefit estimate requires close study of the potential savings offered by the sea-level management option as compared to the money that would otherwise be invested in shoreline protection, the evacuation of populations and the abandonment of 'hopeless cases'. Note that there has already been considerable (if inadvertent) capital investment in sea-level management projects.

Major river diversion and storage will certainly have an impact on ocean dynamics and chemistry²¹. The potential oceanographic alterations are difficult to assess and the ecological consequences of our suggested hydrological cycle interruptions are largely unknown.

The 107 largest freshwater reservoirs already (or will before the end of this century) impound nearly 3,100 km³ of fresh water (Fig. 1), and probably a similar quantity is already stored in the many smaller reservoirs. All this suggests that ~6,000 km³ of water has been temporarily stored in surface reservoirs and prevented from flowing into the world ocean. This would be equivalent to 15.0 mm of sea-level rise or a 12-year lag in sea-level rise at 1.25 mm yr⁻¹. In addition, irrigation projects have absorbed by infiltration ~125 km³ yr⁻¹ since 1932, corresponding to nearly 7,000 km³ of subsurface water

and equal to about 14 years of sea-level rise. Thus reservoir and irrigation storage have caused sea-level rise to lag about 26 years behind its projected uncontrolled rate of rise (Fig. 1). Mankind has thus unwittingly been exercising an appreciable control on sea-level rise. For the past 25 years, continental storage might account for a reduction of as much as 0.75 mm yr⁻¹ in eustatic sea-level rise.

For the immediate future, human ingenuity can undoubtedly find additional capacity to expand storage to accommodate 50 years of sea-level rise. An additional 5 mm of sea-level equivalent could be budgeted to correspond to the increased absolute humidity due to evaporation from all these reservoirs. We project that mankind will exhaust cost-effective water storage capacity near the middle of the next century, thus requiring the emergence of some other means of sea-level-rise management technology.

Although the unintentional sea-level management projects already completed and underway are indeed impressive, further sea-level management will require international cooperation on an unprecedented level. To prevent worldwide littoral disaster, new projects must be planned immediately, and be well underway by the end of this century. All littoral nations must work out some system of contributing to a fund to support these projects. Sea level rise joins 'death and taxes' as the inexorable fate of mankind. This 'fact of life' must be addressed and soon!

Dr Albert kleine Bornhorst and Dr Michael Siebenhuner of Hannover, and Professor Dov Nir of Jerusalem provided helpful guidance concerning the Qattara and Dead Sea Projects. Dr David H. Speidel provided valuable data and welcome direction to sources of information. Earlier versions of this paper were read by David H. Speidel, Michael R. Rampino, Les Marcus, Sergej Lebedeff, Shaik Bornstein, Ann Kelly and Leonard J. Cinquemani. Considerable input came from our colleagues at the Goddard Institute for Space Studies where W.S.N. was a National Research Council Visiting Senior Research Associate in 1980-81 and has since continued his affiliation. Shaik Bornstein programmed and executed the figure. The research was supported by the Research Foundation of the City University of New York. This paper is a contribution to IGCP-200.

Walter S. Newman is at the Department of Geology, Queens College of the City University of New York, Flushing, New York 11367, USA. Rhodes W. Fairbridge is at the Department of Geological Sciences, Columbia University, New York, New York 10027, USA.

1. Gornitz, V. S., Lebedeff, S. & Hansen, J. *Science* **215**, 1611-1614 (1982).

2. Barnett, T. P. *Climatic Change* **5**, 15-85 (1983).

3. IAPSO Advisory Committee on Tides and Mean Sea Level *Eos* **66**, 754-756 (1985).
4. Fairbridge, R. W. & Krebs, D. A. *Geophys. J. R. astr. Soc.* **6**, 532-545 (1962).
5. Meier, M. F. *Science* **266**, 418-421 (1984).
6. Nierenberg, W. A. et al. eds *Changing Climate: Report of the Carbon Dioxide Assessment Committee, 1983* (National Academy Press, Washington, 1983).
7. Barth, M. C. & Titus, J. G. (eds) *Greenhouse Effect and Sea Level Rise* (Van Nostrand Reinhold, New York, 1984).
8. Hoffman, J. S. in *Greenhouse Effect and Sea Level Rise* (eds Barth, M. C. & Titus, J. G.) 79-103 (Van Nostrand Reinhold, New York, 1984).
9. L'vovich, M. I. *World Water Resources and their Future* (American Geophysical Union, Washington, 1979).
10. Tilsley, J. M. *Major Dams, Reservoirs and Hydroelectric Plants—Worldwide* (U.S. Bureau of Reclamation, Engineers and Research Center, Denver, 1983).
11. Postel, S. in *State of the World: 1985* (ed. Brown, L. R.) 42-72 (Norton, New York, 1985).
12. Korzun, V. I. (ed.) *World Water Balance and Resources of the World* (UNESCO, Moscow, 1978).
13. Nace, R. L. in *Water, Earth, and Man* (ed. Chorley, R. J.) 285-294 (Methuen, London, 1969).
14. Ambroggi, R. P. *Scient. Am.* **230**, (5), 21-27 (1977).
15. Steering Committee for the Projected Inter-seas Water Conduit *Interim Report* Pts 1, 2 (State of Israel, Ministries of Energy and Infrastructure and Finance, 1980).
16. Joint Venture Quattara *Quattara Depression Study Feasibility Report* (Lahmeyer International GmbH, Frankfurt, 1981).
17. Micklin, P. P. *Can. Geogr.* **13**, 199-215 (1969).
18. Kelly, P. M., Campbell, D. A., Micklin, P. P. & Tarrant, J. R. *Geojournal* **7** (3), 201-214 (1983).
19. Rich, V. *Nature* **317**, 758 (1985).
20. The Ralph Parsons Co., North American Water and Power Alliance, Brochure 606-2934-19 (1967).
21. Holt, T., Kelly, P. M. & Cherry, B. S. B. *Ann. Glaciol.* **5**, 61-85 (1984).
22. Gornitz, V. S. & Lebedeff, S. *Spec. Publs. Soc. econ. Palaeont. Miner.* (in the press).

REVIEW ARTICLE

Transition region of the Earth's upper mantle

Don L. Anderson* & Jay D. Bass†

* Seismological Laboratory, California Institute of Technology, Pasadena, California 91125, USA

† Department of Geology, University of Illinois, Urbana, Illinois 61801, USA

The transition region of the Earth's upper mantle, 400–650 km deep, appears to be mineralogically and chemically distinct from both the shallow mantle and lower mantle. It contains most of the basalt fraction of the Earth's mantle.

“The transition layer appears to hold the key to a number of major geophysical problems” (ref. 1).

THE Earth's mantle is conventionally divided into three major subdivisions: the shallow mantle, the transition region and the lower mantle. High seismic velocity gradients extend from 400 km, the top of the transition region, to ~800 km. The relatively homogeneous part of the lower mantle starts near 800 km and extends to within several hundred kilometres of the core-mantle boundary. Major changes in mantle mineralogy occur near 400 and 650 km (refs 2–6). The question of whether these represent equilibrium phase boundaries in a homogeneous mantle^{2,6,7} or changes in chemistry^{8–13} is fundamental to many problems in earth sciences. There is abundant evidence that the Earth as a whole is a differentiated body and is inhomogeneous on many scales. For example, the extreme concentration of trace elements in the continental crust requires efficient extraction of these elements from all or most of the mantle^{14,15}. Nonetheless, the chemistry of the mantle remains a controversial issue which can be reduced to three basic questions: (1) is the mantle homogeneous in composition or is it chemically stratified? (2)

Is the major-element chemistry of the mantle more similar to upper mantle peridotites or to chondrites? (3) What is the composition of the source region of basalts erupted at mid-ocean ridges? We address each of these questions using data from cosmochemistry, geochemistry, petrology, seismology and mineral physics.

Chondritic Earth

There is no *a priori* reason to believe that the major and refractory elements were not accreted into the Earth in cosmic (or solar) proportions. But several complicating factors need to be considered; for instance, meteorites and, presumably, the terrestrial planets are depleted to varying degrees in volatile elements such as Na, K, Rb and S. In igneous processes the most refractory crystals are MgSiO₃ (Ca-poor pyroxene) and (Mg, Fe)₂SiO₄ (olivine) and the more fusible (basaltic) components are rich in CaO, Al₂O₃ and TiO₂ and the incompatible trace elements, both refractory and volatile. Melting and melt separation fractionate the major elements, separate peridotite (olivine plus pyroxene) from basalt (Ca and Na-rich pyroxene plus an Al-rich phase), and concentrate the incompatible elements into the liquid fractions. Consequently, extensive melting during accretion separates materials having different crystallization temperatures and densities.

As it turns out, the assumption that the Earth has chondritic abundances of the major rock- and core-forming elements, with oxygen as the light element in the core, provides an excellent fit to the first-order parameters of the Earth such as mean density and size and density of the core (Table 1). The predicted size of the core is remarkably close to the actual size (32 wt%). The composition Fe₂O for the core not only has the appropriate density¹⁶ but also represents the amount of FeO that can be dissolved in molten Fe at core pressures^{17,18}. The core will be slightly larger if it contains chondritic abundances of Ni and some S or if the mantle is less FeO-rich than assumed. Fe₂O, being approximately the eutectic composition of melt in the Fe–FeO system at high pressure¹⁸, has a low melting point and drains to the core. This exercise supports the concept of a chondritic Earth for the major elements, which implies a pyroxene-rich mantle.

Table 1 Estimates of cosmic abundances of the major refractory elements (expressed as wt% oxides) and estimates of the average composition of the subdivisions of the mantle

	Cosmic abundances ⁹⁸ mantle	Shallow mantle (average lherzolite) ¹⁹	Transition region (piclogite) ²²	Lower mantle (cosmic- 20% lherzolite- 10% piclogite)
MgO	36.6	42.2	24.0	36.8
Al ₂ O ₃	3.67	2.05	8.6	3.4
SiO ₂	50.8	44.2	47.0	53.2
CaO	2.89	1.92	8.0	2.4
FeO	6.08	8.29	10.8	4.8
Core				
Fe ₂ O	30.1			

The mantle Fe/Mg ratio is assumed for cosmic abundances⁹⁸.

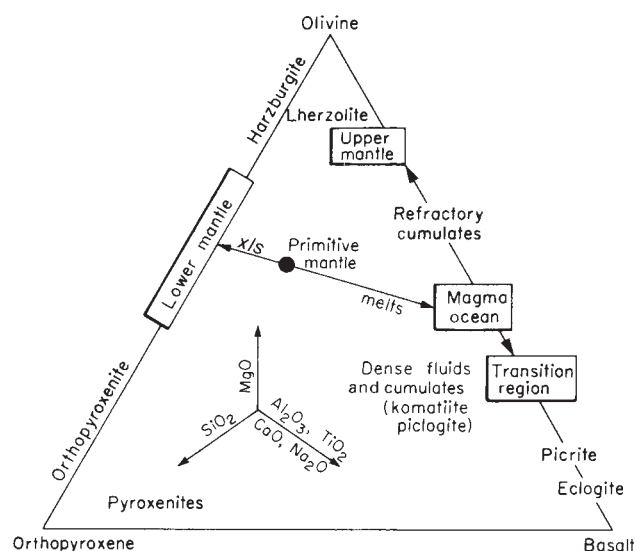


Fig. 1 Schematic illustration of differentiation of primitive mantle (chondritic) during accretion (low-pressure melting) into melts (magma ocean) and refractory residue (*x/s*). Crystallization of the magma ocean separates minerals according to their crystallization temperature and density, giving a chemically stratified upper mantle.

Figure 1 illustrates the accretional differentiation of a chondritic Earth. Gravitational energy is delivered to the surface of a planet and melting and melt-crystal separation are, at least initially, low-pressure processes. A melt layer, or magma ocean, crystallizes refractory, low-density olivine, which is concentrated in the shallow mantle, in peridotite and lherzolite. The dense mineral garnet, $(\text{Mg, Fe, Ca})_3(\text{Al}_2, \text{MgSi, FeSi, CaSi})\text{Si}_3\text{O}_{12}$, crystallizes later and sinks. Residual fluid freezes to a clinopyroxene-garnet-rich mixture with some olivine, an assemblage we term piclogite.

Figure 2 gives the chemistry of a 'cosmic' mantle and several terrestrial rock types. The fields PM are estimates of upper mantle chemistry^{7,19,20} including pyrolite, a hypothetical mixture of basalt and peridotite⁷. Piclogite is the inferred composition of the transition region from seismic data^{21,22}. Most mantle rocks and magmas fall in the garnet-clinopyroxene-harzburgite (mainly olivine) field. Lherzolite and cosmic compositions have an orthopyroxene component. Piclogite falls approximately on the tie line between mid-ocean ridge basalts (MORBs) (and picrite) and harzburgite, thus satisfying a standard assumption about the basalt source region^{7,23}; that is, the source of MORB contains both basaltic and refractory fractions. Piclogite is close in composition to komatiite, a common Precambrian rock type.

Primitive mantle, with cosmic ratios of refractory elements, contains the equivalent of 50% peridotite, 36% orthopyroxene and 9–14% MORB (Table 2). If basalts form by melting in the shallow mantle, where melt segregation can occur at relatively low degrees of melting, the MORB reservoir may contain only a small amount of basalt, ~20%. If this reservoir is deeper than 200 km, melt segregation is more difficult^{24,25} and extensive melting can occur, resulting in instability of the source region, adiabatic ascent and extensive melting before eruption at mid-ocean ridges or intrusion into the shallow mantle. In this case basalt may represent a much larger fraction of the reservoir. The transition region, which we argue later is MORB plus up to 30% olivine, represents 11% of the mantle. The basalt fraction of the mantle may not be uniformly dispersed because the transition region itself could accommodate most of the basalt of the Earth.

Basaltic melts are buoyant at low pressure but are denser than olivine and pyroxene below some 200 km (ref. 24). Crystallization of basalt at high pressure yields dense eclogite. The original

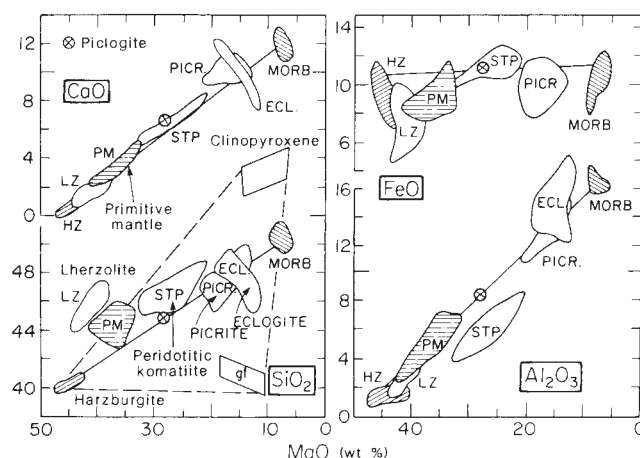


Fig. 2 Chemistry of primitive mantle and various products of mantle differentiation. Data from refs 15, 19, 20, 22. Piclogite, similar in composition to komatiites, satisfies the seismic data for the transition region²². Picrite may be the parent magma for MORBs (ref. 23). The lower mantle is the result of removal of lherzolite and picrite from cosmic abundances. The fields labelled PM refer to petrological models of the upper mantle (refs 7, 19, 20). Lherzolite (LZ) and harzburgite (HZ) are representative of shallow mantle peridotites. PICR., picrite; ECL., eclogite; STP, spinifex-textured peridotites. The shaded areas represent the most abundant rock types. Dashed lines enclose the HZ, Cpx, garnet field. The oxides (ordinate) are plotted against MgO (abscissa).

stratigraphy of the mantle would therefore be, from the top down, basalt, peridotite, eclogite and, at great depth, olivine orthopyroxenite, the high-pressure residual of the original differentiation. This is a gravitationally stable configuration except that when eclogite partially melts it becomes buoyant because of the elimination of garnet.

If the upper 400 km of the mantle is lherzolite¹⁹ and the transition region is piclogite^{21,22}, the composition of the lower mantle (Table 1) is considerably enriched in Si compared with lherzolite. The lower mantle in this model is 85 wt% MgSiO_3 and 7.4 wt% $(\text{Mg, Fe})\text{O}$. Upper mantle lherzolite, by contrast, has 22.2 wt% $(\text{Mg, Fe})\text{O}$.

The measured Mg/Si ratio of peridotite is higher than in any meteorite class. This has been attributed to volatility and net depletion of the mantle in Si (ref. 26). However, basalt extraction causes an increase in Mg/Si and a decrease in density of the residue; this residue will therefore concentrate in the shallow mantle^{27–30}. Ca–Al-rich chondrites are depleted in volatiles but still have a lower Mg/Si ratio than lherzolites. There is no reason to suppose that the Earth's mantle is depleted in Si but several reasons to suspect that upper mantle lherzolites have been depleted in Si by removal of basalt. Chondritic ratios of Ca/Al and Al/Mg limit the amount of basaltic material. Orthopyroxene, a possibly important constituent of the lower mantle, is required to achieve chondritic ratios of Mg/Si.

The mantle + crust is depleted in volatile elements relative to carbonaceous chondrites^{14,15,31}. Assuming that refractory elements occur with cosmic ratios, it is possible to estimate the

Table 2 Components of primitive mantle (wt%) that yield chondritic ratios of refractory lithophiles

	Model 1	Model 2
Continental crust	0.55	0.55
Peridotite	52.2	49.2
Orthopyroxenite	38.0	35.4
Basalt (MORB)	9.3	14.4
Kimberlite	0.19	0.17

Composition of components from refs 15, 19, 79.

Table 3 Composition of primitive mantle based on multi-component petrological models

	Cosmic	Mantle		% Of element in:			
		Model 1	Model 2	Crust	Model 1 MORB	Model 2 Crust	Model 2 MORB
Mg (%)	21.3	21.1	21.4	0.1	2.6	0.1	4
Al (%)	1.9	1.9	1.9	2.8	46	2.4	69
Si (%)	23.1	22.9	23.3	0.6	9.5	0.7	15
Ca (%)	2.0	2.0	2.0	1.5	43	0.9	67
Ti (%)	0.1	0.08	0.12	3.4	56	1.7	58
Sr (p.p.m.)	17.3	13.3	13.4	17	36	21	56
Ba (p.p.m.)	4.9	6.2	8.0	31	6	49	7
La (p.p.m.)	0.51	0.60	0.57	18	16	27	26
Yb (p.p.m.)	0.35	0.37	0.26	3	42	3	93
Th (p.p.m.)	0.06	0.064	0.051	41	4	55	8
U (p.p.m.)	0.02	0.019	0.015	35	9	47	19
Li (p.p.m.)	3.4	2.0	2.6	3	56	2	67
Na (%)	1.1	0.12	0.13	12	33	13	49
K (%)	0.12	0.035	0.017	19	8	56	27
Rb (p.p.m.)	4.97	0.41	0.50	56	4	68	5
Cs (p.p.m.)	0.41	0.02	0.01	62	1	67	2
Ratios							
Mg/Si	0.92	0.92	0.92				
Ca/Al	1.07	1.07	1.03				
Mg/Al	11.3	11.4	11.1				
La/Yb	1.48	1.61	2.19				
Ce/Yb	3.87	3.43	4.31				
Rb/Sr	0.288	0.031	0.037				

p.p.m., parts per 10⁶.

major- and trace-element chemistry of the mantle from various observed products of mantle differentiation¹⁵. Attempts to do this with three-component systems (crust, basalt and peridotite for major elements³²; or crust, depleted mantle and primitive mantle in isotope studies^{33,34}) fail to give chondritic ratios for all refractory lithophile elements or for key major-element ratios as Mg/Si. To span the trace-element and isotope chemistry of basalts and the major-element chemistry of ultramafic rocks, it is necessary to involve at least four components¹⁵. The crust is also an essential component of primitive mantle. We therefore treat primitive mantle as a five-component system (crust, basalt (MORB), enriched magma (kimberlite or ocean-island basalt), lherzolite and pyroxenite) and solve for the mixing ratios which give chondritic ratios for the refractory lithophile elements. Typical solutions are given in Table 2. The peridotite component varies from ~47 to 60% depending on its fertility (that is, Al₂O₃ and CaO content). The enriched component is ~0.2% if it is similar to kimberlite and somewhat more if it is similar to ocean-island basalts or recycled crust. The MORB content of primitive mantle ranges up to 14 wt%, the amount depending, again, on the fertility (or basalt content) of the peridotite.

The continental crust contains >30% of the large and high-valence ions (Table 3). MORB is complementary to continental crust, being depleted in the very incompatible elements and enriched in Ca, Al, Hf, Yb, Y and other elements that substitute for Al and Ca in garnet. The enriched mantle component, required because of the enriched nature of ocean-island basalts³⁵, xenoliths³⁶ and kimberlites^{37,38}, may represent subducted crust or material left over from inefficient melt removal from the mantle. We use kimberlite to model this material but, for most elements, continental crust or ocean-island basalts would serve as well. Table 3 gives results from two mass-balance calculations, including the fraction of various elements contained in the crust and in the basalt part of the MORB reservoir. The extreme enrichment of some elements in the continental crust provides strong evidence against a substantial primitive, undifferentiated reservoir in the mantle. The whole mantle must be processed to obtain these enrichments; most of the mantle must therefore be depleted in these elements. The most efficient way to accomplish this separation is during the melting and vaporization associated with accretion. Stripping incompatible

elements out of the mantle by subsequent melting is much less efficient—the last fraction of melt is difficult to remove. The concentration of W, Cs, Rb and Th in the continental crust is such that >50% of the mantle must have been completely stripped of these elements. The MORB source contains a large fraction of the mantle's inventory of Al, Ca, Ti, Sr, Li, Na and the heavy rare-earth elements (HREE) and high concentrations of the light rare-earth elements (LREE), Sr, Y, Zr, Hf, Na, Li, U and Ta. These elements are presumably sequestered at depth in clinopyroxene and garnet, the major minerals of eclogite. The continental crust is excessively enriched in elements with large ionic radius or high charge, whereas MORB is enriched in elements with ionic radii between Al³⁺ and Ca²⁺.

These results strongly suggest an extensively differentiated Earth, probably involving a deep magma ocean in early Earth history. In contrast to the Moon, the terrestrial crust is trivial in volume. It does not contain most of the Ca²⁺ and Al³⁺ of the planet and it is deficient in elements which can be removed from a melt by the crystallization of garnet and clinopyroxene. MORBs are enriched in these elements, suggesting that they are melts from a deep eclogite-rich layer, the terrestrial equivalent of the thick lunar crust. The fact that the roots of mid-ocean ridges can be traced by seismic techniques to depths of the order of 400 km³⁹⁻⁴² and that the transition region appears to be garnet-rich rather than olivine-rich^{21,22,43}, support the notion that the transition region is the major basalt reservoir. The transition region would then contain much of the mantle inventory of the HREE, Al, Ca, Ti, Sr, Li and Na, leaving a lower mantle depleted in Ca, Al, Ti and incompatible elements including U, Th and K. If the shallow mantle is predominantly residual or cumulate olivine, the most refractory mantle mineral at low pressure, the lower mantle will be even more pyroxene-rich than primitive mantle. Trace element-modelling (including rare-earth elements, REE) of MORB genesis requires that most of the garnet and clinopyroxene in the source be consumed by the melting process²⁷⁻³⁰.

Upper mantle and 400-km discontinuity

The upper mantle includes the seismic lithosphere, which averages ~40 km in thickness under ocean basins⁴⁴ and 150 km under stable shields⁴⁵. The shield lithosphere and the top of the

oceanic lithosphere have seismic properties consistent with peridotite⁴⁶. The lower oceanic lithosphere may be garnet- rather than olivine-rich^{8,46}. The 150-km-thick shield lithosphere is probably buoyant with respect to oceanic lithosphere and the underlying mantle, which explains its long-term stability⁴⁶.

Seismic velocities decrease at 20–50 km under oceans and 150 km under shields, indicating a change in mineralogy and, for some regions, the presence of a melt phase. Seismic velocities are so low in some regions that partial melting is implied to depths at least as great as 300 km (ref. 46). Under some mid-ocean ridges low velocities extend to 400 km (refs 39–42).

Seismic velocities increase abruptly at 400 km, which is usually attributed to the olivine-spinel phase change. However, the velocity jump associated with this transformation^{47,48} is ~20% and is spread over an appreciable depth interval^{5,49,50} whereas the observed jump is 4–5% and is rather abrupt. The other important minerals of the mantle, pyroxenes, transform to garnet solid-solution assemblages (garnetite), primarily between 350–400 km for orthopyroxene^{50–52} and somewhat greater depth for clinopyroxene⁵³. The associated velocity increase is greater than for the olivine-spinel transformation^{61,62}; this places an upper bound on the olivine plus orthopyroxene content of the mantle near 400 km. Garnet and possibly jadeite ($\text{NaAlSi}_2\text{O}_6$) survive to high pressures⁵⁴, dilute the effects of the other minerals and reduce the size of the discontinuity. If the mantle above 400 km is mostly olivine and orthopyroxene, the 400-km discontinuity must involve a change in chemistry.

Transition region

In early seismological models the transition region was characterized by a rapid but smooth increase in seismic velocities between 400 and 1,000 km depth. As the quality of seismic data improved, the transition region was found to be more complex than originally thought. Abrupt changes in velocity were found at ~400 and 650 km (refs 55–60). Recent models^{41,42,45} have velocity jumps of 4–6% at each discontinuity, high gradients between the discontinuities and below the deeper one to ~750 or 800 km (refs 41, 42, 44).

The major minerals in shallow mantle peridotites are olivine $[(\text{Mg}, \text{Fe})_2\text{SiO}_4]$ and orthopyroxene $[(\text{Mg}, \text{Fe})\text{SiO}_3]$. Olivine collapses to the β -spinel structure at ~400 km, with a large increase in density and elastic properties, and to γ -spinel at ~500 km, with a small change in physical properties⁴⁷. Phases with lower seismic velocities are required to satisfy the seismic constraints. Garnet and clinopyroxene, a solid solution of diopside $[\text{Ca}(\text{Mg}, \text{Fe})\text{Si}_2\text{O}_6]$ and jadeite, have the appropriate characteristics, but these are minor minerals in peridotite. Orthopyroxene transforms to a garnet-like phase, majorite, starting at ~300 km. The modulus of majorite⁶¹ is much higher than garnet or spinel and the modulus in the transition region^{21,22,62} would exclude a spinel-majorite mineralogy (but see ref. 100).

The seismic properties of β - and γ -spinel are similar^{47,48} and only a small discontinuity would result from the β - γ transformation. The orthopyroxene-majorite transition is complete at depths of the order of 400 km (refs 50–53). The only further transformation expected at transition-zone pressures is the garnetite-to-ilmenite structure phase change, which probably gives a slight increase in velocity. Aside from a change in chemistry the only way to explain the low velocities and high gradients in the transition zone is that clinopyroxene transforms to a Ca-rich garnet-like phase analogous to majorite. This transition is spread out over a broad pressure interval and is much deeper than the corresponding transition in orthopyroxene⁵³. Thus, the velocities, the velocity gradient and the great depth of the mixed-phase region all argue for a clinopyroxene-garnet-rich transition region, with the clinopyroxene gradually changing to a Ca- and Si-rich garnet. Whether the above inferences are plausible depends on the high-pressure stability of clinopyroxene and garnet.

There have been several attempts to model the seismic

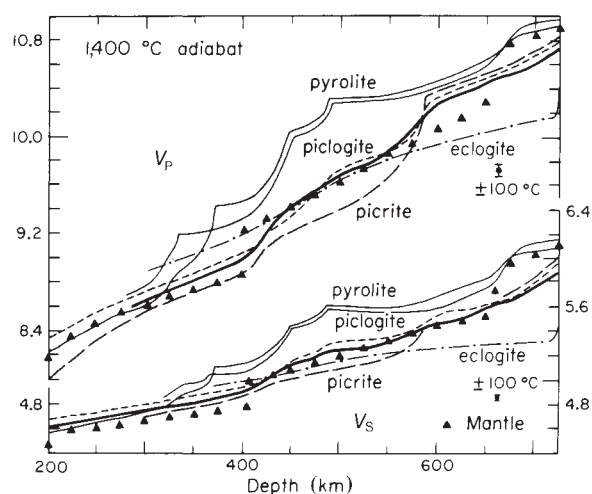


Fig. 3 Compressional (V_p) and shear (V_s) velocity as a function of depth for PREM (triangles⁹⁵) and for various petrological models, using finite strain theory⁶⁴ and parameters summarized in refs 22 and 50. Phase relations modified from refs 5, 50, 52, 53, including the pressure correction of ref. 50. The rapid increases in velocity are caused primarily by orthopyroxene-majorite, olivine-spinel, clinopyroxene-majorite, spinel to perovskite plus magnesio-wustite and garnet-majorite to perovskite. The garnet-ilmenite transition is ignored. Piclogite contains more olivine and less garnet than picrite and eclogite. Heavy solid lines, piclogite; thin solid lines, pyrolite; long dashed lines, picrite; dot-dashed lines, eclogite; short dashed lines, a garnet pyroxenite. If the new K_s value for majorite (ref. 100) is used instead of the Jeanloz value (ref. 61) then the velocities in pyrolite are lower but the shape and gradient are still wrong. PREM, preliminary reference Earth model.

velocities in the transition region for olivine- and orthopyroxene-rich (for example, peridotite) compositions^{21,62} using mineral physics and phase-equilibrium data^{2,3,10,21,22,62}. The velocities, gradients and discontinuities are not well matched by homogeneous peridotite models. Using current estimates of appropriate phase diagrams^{22,49–53} and finite strain theory^{63,64} we have calculated seismic velocities for several mineralogical assemblages (Fig. 3). Picrite and eclogite are garnet- and clinopyroxene-rich rocks with different garnet/clinopyroxene ratios. The composition of piclogite is given in Table 1 and Fig. 2. It is clear that the velocities in pyrolite are higher than observed. The heavy solid line, piclogite (Fig. 3), results from an attempt to satisfy the velocities between 200 and 650 km by a single, homogeneous mineralogy. Note that the 400- and 650-km discontinuities are not well matched. It seems therefore that neither of them can be modelled as equilibrium phase boundaries. Eclogite and piclogite, with appropriate garnet/clinopyroxene ratios, satisfy the data between 400 and 550 km. In the eclogite model the 400-km discontinuity would be a chemical change from peridotite to eclogite. In the piclogite model the discontinuity is mostly caused by the olivine- β transition. Piclogite contains 30 wt% olivine.

Weidner⁶⁵ calculated seismic velocities in the transition region and his results apparently support the pyrolite model for depths between 400 and 500 km. He assumes, however, that majorite has the same shear modulus, G , as garnet despite the fact that the bulk moduli, K_s , differ by >20%. Our estimate of K_s/G for majorite, 1.86, is similar to other garnets and in the typical range for Fe-poor silicates. Weidner's value, 24% greater, is representative only of Fe-, Co-, Cd- and Ge-rich compounds, but these have very low K_s compared with majorite. Fe-poor and high- K_s oxides and silicates all have $K_s/G < 1.9$. Weidner's model also has appreciable orthopyroxene, a low-velocity phase, present to depths as great as 500 km, whereas orthopyroxene has transformed to majorite by 400 km in our calculations. Below 500 km the two studies agree in that the compressional velocity

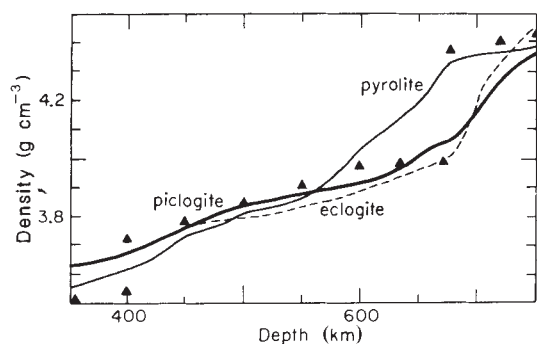


Fig. 4 Density versus depth calculated for various petrological models along the 1,400 °C adiabat. Note that piclogite and eclogite are gravitationally stable in the transition region and olivine-rich assemblages, such as pyrolite, are stable in the shallow mantle. $\blacktriangle$, PREM. Heavy solid line, piclogite; thin solid line, pyrolite; dashed line, eclogite.

for pyrolite exceeds the seismic values. Even if majorite has elastic properties similar to garnet¹⁰⁰ the velocities in pyrolite would be too high below 400 km.

It is interesting to compare the density profiles for pyrolite, piclogite and eclogite (Fig. 4). Eclogite and piclogite have low densities to ~670 km and approach lower mantle values near 750 km. The seismic resolution of density is poor but the strong reflector near 650 km implies an abrupt jump in density as well as in velocity. None of the isochemical models has this characteristic, but a chemical change from eclogite or piclogite to MgSiO_3 -perovskite would have the correct magnitude of jump. The density jump in a pyrolite mantle is large but diffuse, and the changes in elastic properties are also probably diffuse. Therefore, phase changes in pyrolite do not explain the 650-km discontinuity. Note that piclogite and eclogite are less dense than pyrolite and the mantle between 670 and 750 km. Eclogite-rich assemblages will therefore be unable to sink into the lower mantle. Ringwood⁶⁶ reached the opposite conclusion because of the large amount of stishovite (a dense polymorph of SiO_2) in his quartz eclogite and because he assumes that CaSiO_3 forms perovskite, rather than garnet-majorite, and that it is denser than the lower mantle. It is more likely that CaSiO_3 is present in diopside or garnet rather than as an isolated phase. Refractory peridotite is less dense than fertile peridotite or piclogite and will concentrate in the shallow mantle. The bulk of the material subducting into the transition region may be eclogitic, possibly an important constituent of the lower oceanic lithosphere.

The main disagreement between piclogite and seismic data is the increase of the velocity anomaly V_p in piclogite at a depth of 550 km caused by the garnet-perovskite phase change. It is probable that garnet transforms to perovskite via an intervening field of lower-velocity ilmenite. A thermal boundary layer at 650 km combined with a positive Clapeyron slope would also increase the depth of this phase change. The mineralogy implied by the seismic velocities can be used to calculate the density contrasts across the 400- and 650-km transitions. These are 3 and 10%, respectively. The effect of temperature, including phase changes and partial melting, can reverse the density contrast at 400 km but is unlikely to be able to do so at 650 km. A chemical change at 650 km, perhaps depressed to 720 km under subduction zones⁶⁷ and elevated to 620 km in warmer parts of the mantle^{68,69} is consistent with the abrupt termination of seismic activity⁷⁰, the focal mechanisms of deep earthquakes^{71,72} and the deep resistance encountered by slabs⁷³.

Lower mantle mineralogy

The 650-km discontinuity was discovered and characterized⁵⁵⁻⁶⁰ in the 1960s and is therefore one of the most recently mapped major features of the mantle. In the context of an olivine-rich mantle the discontinuity was interpreted as a phase change from

the spinel form of olivine to a hypothetical post-spinel phase, with silicon in octahedral coordination and an increase of FeO content^{2,3}. Thus, from the beginning, there has been an element of both a phase change and a chemical change associated with the interpretation of this boundary. Uncertainties in crystal structures and elastic properties of high-pressure phases, however, made it difficult to exclude a purely equilibrium-phase boundary. This interpretation required a strongly negative Clapeyron slope for the reaction² and a particularly high density and bulk modulus for the post-spinel phase^{3,7}. Later work showed that the density and elastic properties of the lower mantle were higher than the post-spinel phase. A lower mantle stoichiometry closer to pyroxene than to olivine, and closer to chondritic than to peridotitic, was proposed^{3,8-13,63}. Although it is difficult to estimate the FeO content of the lower mantle^{6,8,74}, variations in the SiO_2 content have large effects on the seismic velocities¹¹, whether the SiO_2 is in stishovite or perovskite^{8,11-13}.

The 650-km discontinuity is a good reflector of short-period seismic energy. To be a good reflector requires a large increase in density and velocity over a small fraction of a wavelength, perhaps 3-4 km (refs 62, 68, 75, 76). Most phase changes in mantle silicates are second order and are spread over 20 km or more⁵⁰. Reflections and conversions of seismic waves have been reported from depths ranging from ~590 to 720 km (refs 67-69, 75). Thus, there is a possible range of 130 km in the depth of the 650-km discontinuity.

Earthquake activity terminates abruptly near 670 km. Detailed studies indicate that the Tonga-Fiji slab encounters a barrier at 670 km, then shortens and thickens⁷¹. The transition to horizontal shear in the deepest part of the seismic zone is evidence of the strong resistance experienced by subducted material at the discontinuity. Material appears to move laterally rather than vertically at the base of the seismic zone⁷¹. High velocities below some deep-focus earthquakes in the Kuriles⁷⁷ could be caused by anisotropy, with high velocities in the flow plane, or to deformation of the interface (ref. 78; B. Marsh and B. Hager, personal communication) and detachment of a boundary layer at the top of the lower mantle.

In a cold slab the stable form of orthopyroxene below some 570 km is the ilmenite rather than the garnet structure⁷⁹. Ilmenite is remarkably anisotropic⁸⁰ (16% for V_p , 34% for V_s). If it crystallizes with its principle axes in the plane of the slab, then seismic waves travelling down dip and along strike will be very fast compared with waves that travel along the slow intermediate directions or exit the slab early to enter the warmer adjacent mantle with lower velocity and more isotropic phase assemblages. This would give residuals having similar characteristics to a deeply penetrating slab. Craeger and Jordan⁷⁷ use observations on olivine-rich rocks to dismiss the importance of anisotropy, but olivine, of course, is not stable at the required depths. Ilmenite differs from olivine in the manner required to explain the observations, in particular in the magnitude and orientation of the S -anisotropy, without requiring slab penetration into the lower mantle. The deep slab penetration model⁷⁷ requires a large (5%) velocity anomaly in the slab at great depth, much larger than found at shallower depths⁸¹. Actually, slab material may be less dense and have slower seismic velocities than the lower mantle (Fig. 4), particularly if ilmenite is present. The high gradient of seismic velocity between 650 and 750 or 800 km implies a high density gradient. Slabs may therefore penetrate ~100 km into the lower mantle before they become buoyant. A similar amount of depression of the boundary by the cold slab is also possible. Crumpling of the slab at 650 km (ref. 71) will destroy its slab-like character even if it can sink into the lower mantle. The correlation of geoid highs with subduction regions also indicates a strong resistance of slab material to penetration into the lower mantle⁷³. The 650-km discontinuity may be composed of two reflectors 40 km apart⁶⁹, which may be further evidence for a slab lying on the discontinuity^{71,72}.

There is general agreement that a change in mineralogy (that is, change in crystal structure or phase) occurs near 650 km. Early thermochemical models viewed the discontinuity as an equilibrium phase boundary between olivine in the spinel structure and a homogeneous post-spinel phase having the properties of the constituent oxides. The nature of this post-spinel phase was unknown. Analogue and shock-wave data ultimately lead to the proposal that (Mg, Fe)SiO₃ (perovskite) and (Mg, Fe)O are the dominant phases⁸²⁻⁸⁴. Perovskite is stable at depths >750 km and at lower pressure if the Clapeyron slope is negative⁸²⁻⁹⁴. Perovskite was predicted to have a slightly higher bulk modulus than the mixed oxides⁸⁻¹¹ and this is confirmed by recent measurements⁸⁵. The lower mantle, below 800 km, had velocities in excess of those predicted for peridotite and conclusions regarding a pyroxene-rich mantle were supported^{8,11}. The conclusion that the lower mantle is pyroxene-rich is independent of the mineralogy, that is, perovskite + (Mg, Fe)O or (Mg, Fe)O + stishovite. The uncertainties in the elastic properties of the two assemblages overlap.

Three methods have been used to infer lower mantle composition: (1) extrapolation of laboratory data to lower mantle conditions^{9,10}; (2) comparison of seismic and shock-wave data^{8,12}; and (3) extrapolation of seismic data to standard conditions^{3,11,63}. The last method is the most accurate and involves the fewest assumptions. The seismic data themselves are used to estimate density, elastic properties and pressure derivatives and to test for adiabaticity and homogeneity.

The seismic parameter $\Phi = K_s/\rho = V_p^2 - (4/3)V_s^2$, inferred from seismic, compression and ultrasonic data, is a useful diagnostic of crystal structure and SiO₂ content (K_s , ρ , V_p and V_s are the bulk modulus, density, compressional and shear velocities, respectively.) The variation of Φ with depth in most of the lower mantle is consistent with the assumptions of homogeneity and adiabaticity^{11,95} which are required to extrapolate to pressure (P) = 0 to obtain Φ_0 and ρ_0 . The major minerals of the lower mantle can have similar densities^{12,86} but quite different Φ .

Si coordination and the Si-O bond distance are the same for perovskite and stishovite⁹⁴. Therefore, K_s and Φ of the lower mantle depend on SiO₂ content and are nearly independent of details of the mineralogy. SiO₂, whether in perovskite or stishovite, increases Φ whereas CaO, FeO and MgO decrease it. Tables 4 and 5 give Φ values for possible lower mantle assemblages including lherzolite, undepleted and depleted cosmic mineralogies and pure MgSiO₃ (perovskite). Also given is the seismic Φ corrected for temperature and pressure; Φ_0 (P = 0) is $54 \pm 2 \text{ km}^2 \text{ s}^{-2}$. Assuming an adiabat (temperature T_0 = 1,400–1,800 °C) yields Φ_0 in the range $62 \pm 3 \text{ km}^2 \text{ s}^{-2}$. Olivine, with equal molar proportions of MgSiO₃ (perovskite) and (Mg_{0.8}Fe_{0.2})O, gives $52 \text{ km}^2 \text{ s}^{-2}$. Pure perovskite gives $63 \text{ km}^2 \text{ s}^{-2}$. Clearly, Φ_0 for an olivine or lherzolite lower mantle is too low; the pyroxene and cosmic models, with SiO₂ > 52 wt %, provide a good fit. The differentiated cosmic model is cosmic composition minus a lherzolite shallow mantle and a piclogite transition region. Barren cosmic is calculated by removing all of the basaltic fraction from primitive mantle. The increase in Φ associated with phase changes in lherzolite (γ -spinel + majorite) to the post-spinel phases is too small to be consistent with the observed increase unless the K_s (majorite) of ref. 61 is too high.¹⁰⁰ There is a trade-off between the effects of composition and temperature^{12,50,87}, but the low value of thermal expansion implied for perovskite in a pyrolite model⁸⁷ is unlikely⁸⁸. Thus, the composition of the lower mantle is consistent with a differentiated chondritic Earth model.

Surface-wave tomography

Surface-wave tomography^{39,40,96} has put regional seismic studies into a global context. The low-velocity layer is not a uniform global layer of constant thickness. It rises to the near surface under young oceanic, back-arc and extensional regions and is

Table 4 Seismic parameter Φ at $P = 0$

	Φ (km s^{-1}) ²
Lower mantle	
High temperature	53.4
Corrected for temperature	63 ± 2
Perovskite	62 ± 6
Al ₂ O ₃	63.4
SiO ₂ (stishovite)	73.7
MgO	47.4
(Mg _{0.8} Fe _{0.2})O	40.8

Data from refs 11, 12, 50, 61, 87, 94, 99.

Table 5 Composition and seismic parameter Φ for minerals, mineral assemblages and the lower mantle

	SiO ₂	MgO wt %	Al ₂ O ₃	Φ (km s^{-1}) ²
Transition region minerals				
γ -Spinel	43	57	0	50.3
Majorite	60	40	0	62.5*
Lower mantle assemblages				
lherzolite	44.8	42.8	2.08	55.2
Post-spinel	43	57	0	55.4
Perovskite	60	40	0	53.4
Cosmic				
Undifferentiated	50.8	36.6	3.67	58.3
Differentiated	52.4	38.7	2.64	59.2
Barren	50.4	44.3	0	58.0
Lower mantle observed				63 ± 2

* Ref. 61. Φ = 45.5 from ref. 100.

deep, >150 km, and less pronounced under shields. In many regions low velocities can be traced down to the transition region, although the global pattern below 400 km bears little relation to surface tectonics. Partial melting is implied for the slower regions⁴⁶. The low velocities associated with magmatic centres seem to be broad upwellings originating at depths at least as great as 400 km.

The most pronounced regions of inferred upwelling in and above the transition region are: North Atlantic, Darwin Rise, North-East Africa and South-East Asia. Most of these have low velocities at the top of the lower mantle⁹⁷, lie in geoid highs and have high average elevations. The correlation with velocities at the top of the lower mantle indicates that material could be rising through the 650-km discontinuity or that hot regions in the lower mantle control the locations of hotspots in the upper mantle.

Thus, seismic tomography has provided both the first global view of the structure of the mantle and evidence that it is not a simple layered system. Heterogeneities in velocity and anisotropy imply that magmas are generated from reservoirs deep in the upper mantle or transition region. The recognition that mid-ocean ridges are deeply rooted suggests that entrainment and melting of shallow mantle, and magma mixing, could be associated with the uplift of the MORB source or ascent of its separated magma. The view that emerges is of broad, deep upwellings, some of which are associated with ridges and continental rifts. Oceanic crust and other extrusives probably represent only a small fraction of the molten material delivered to the upper mantle. The rest presumably spreads laterally and is available for off-ridge volcanism and underplating of the lithosphere. The oceanic lithosphere may, in fact, thicken by underplating of material similar to the parent magma of oceanic basalts, giving a dense, eclogite lower lithosphere. The view of a homogeneous partially molten peridotitic asthenosphere extending to ~200 km and passively providing basalts to mid-

ocean ridges clearly must be abandoned. The MORB source is much deeper but the shallow mantle may be the holding tank for large amounts of picritic material and the repository for light differentiation products of all ages, including peridotite.

Discussion

There are no inconsistencies with a chondritic Earth when the major and refractory elements are considered. The silicon deficiency hypothesis is unnecessary. Seismic data are consistent with a differentiated Earth having a peridotite shallow layer, a basalt-rich transition region, a pyroxene-rich lower mantle and a dynamically stratified upper mantle. Isotope data show that mantle inhomogeneities are ancient, possibly developing during the accretion of the Earth and its high-temperature early history. The energy of accretion is adequate to melt or even vaporize the infalling material. A thick early atmosphere and buoyant insulating layers in the upper mantle would delay crystallization of melts trapped at depth and promote gravitational stratification during accretion, cooling and crystallization.

Much of the discussion of whole mantle against layered convection focuses on the roles of viscosity, mean atomic weight and phase changes. These effects alone can probably not prevent whole-mantle convection. On the other hand, an isoviscous mantle with constant mean atomic weight can be chemically stratified. For example, a differentiated planet with a feldspar-rich crust, an olivine-rich shallow mantle, a garnet-rich transition region and a perovskite lower mantle is gravitationally stable. Variations in SiO_2 , MgO and Al_2O_3 , which have similar mean relative molecular mass, control the stable mineral assemblages which, in turn, determine the intrinsic density and seismic properties.

Chemical stratification of the mantle is a natural consequence of melting and melt-solid separation. Basalts and cumulates are examples of these processes; the crust and core are more global manifestations. Convective stirring to maintain or regain homogeneity is a process for which there is little support. A chemically stratified mantle does not imply that each layer is isotopically homogeneous on a global basis. Mixing and homogenization may take place in a given convection cell or between adjacent cells but there is little communication between opposite sides of the Earth. In a dynamically stratified Earth the shallow mantle is enriched at various times by subduction of sediments and altered oceanic crust, and by upward migration of melts. A deeper reservoir may provide the bulk of the material erupting at oceanic islands, mid-ocean ridges and mid-plate environments.

A garnet- and clinopyroxene-rich region is stable at depth when it is cold but is buoyant when it is hot or partially molten. Therefore, it can be responsible for both radial and lateral inhomogeneity in the upper mantle. The broad stability field of garnet-majorite prevents it from converting to perovskite and sinking into the lower mantle. Although the lower mantle may be convecting, it can be isolated chemically, but not thermally, from the upper mantle because of its high intrinsic density. In contrast, the density differences between the shallower layers can be reversed by high temperature, partial melting and phase changes.

Differentiation of the Earth leads to boundaries which separate materials differing in both mineralogy and composition. Basaltic melts rise until they become neutrally buoyant or until trapped by the lithosphere. Subducted basalt and melts intruded at depth convert to eclogite and sink. Phase changes in basalt-eclogite occur at different depths from those in olivine- and orthopyroxene-rich materials. Thus, if the 650-km discontinuity were initially an equilibrium-phase boundary in a homogeneous mantle, subducted eclogite or a piclogite cumulate would not be able to sink through the boundary if the transformation of $\text{CaO-Al}_2\text{O}_3\text{-Na}_2\text{O}$ -rich garnetite to denser phases occurs at greater pressure or results in less-dense minerals than involved in the original equilibrium phase boundary.

We refer to the mantle as dynamically stratified in contrast to the static layering which is the ultimate fate of a cooling planet. When the intrinsic density difference between reservoirs is smaller than temperature-induced variations in density, overturn can occur. These density variations are caused by thermal expansion and by phase changes such as basalt-eclogite, exsolution of pyroxene and partial melting. Thus, piclogite or fertile peridotite is denser than residual peridotite when cold but can be less dense when hot. Adiabatic ascent of the deeper material can cause melting in both layers, entrainment and mixing leading to a lumpy shallow mantle, and magmas which span a large chemical and isotope range. Crystal fractionation of hybrid magmas before eruption induces even more spread in their geochemical characteristics. These phenomena, as well as subduction, create isotope heterogeneities in the shallow mantle. A dynamically stratified mantle and a shallow mantle, heterogeneous on all scales, can coexist. The high density of cold eclogite, and the low density of hot eclogite or basalt, however, means that material of basaltic chemistry will eventually separate from material of peridotite chemistry with only the smaller basaltic accumulations being potentially trapable in the shallow mantle.

We thank B. Hager, E. Stolper, B. Marsh, M. O'Hara and S. Maaloe for stimulating discussions and R. Jeanloz, D. Yuen and D. Weidner for preprints. This work was supported by NSF grants EAR-8115236 and EAR-8317623, and NASA grant NSG-7610. Contribution 4206, Division of Geological and Planetary Sciences, California Institute of Technology.

- Birch, F. J. *geophys. Res.* **57**, 227-286 (1952).
- Anderson, D. L. *Science* **157**, 1165-1173 (1967).
- Anderson, D. L. *Miner. Soc. Am. Spec. Pap.* **3**, 85-93 (1970).
- Ringwood, A. E. *Phys. Earth planet. Inter.* **3**, 109-155 (1970).
- Akimoto, S., Matsui, Y. & Syono, Y. in *The Physics and Chemistry of Minerals and Rocks* (ed. Strens, R.) 327-364 (Wiley, New York, 1976).
- Anderson, D. L. *Geophys. Res. Lett.* **3**, 347-349 (1976).
- Ringwood, A. E. *Composition and Petrology of the Earth's Mantle* (McGraw-Hill, New York, 1975).
- Anderson, D. L. *A. Rev. Earth planet. Sci.* **5**, 179-202 (1977).
- Gaffney, E. S. & Anderson, D. L. *J. geophys. Res.* **78**, 7005-7014 (1973).
- Burdick, L. J. & Anderson, D. L. *J. geophys. Res.* **80**, 1070 (1975).
- Butler, R. & Anderson, D. L. *Phys. Earth planet. Inter.* **17**, 147 (1978).
- Watt, J. P. & Ahrens, T. J. *J. geophys. Res.* **87**, 5631-5644 (1982).
- Sawamoto, H. in *High Pressure Research* (eds Manghnani, M. & Akimoto, S.) 219-244 (Academic, New York, 1977).
- Gast, P. in *The Nature of the Solid Earth* (ed. Robertson, E. C.) 19-40 (McGraw-Hill, New York, 1972).
- Anderson, D. L. *J. geophys. Res.* **88**, B41-B52 (1983).
- Bullen, K. E. *The Earth's Density* (Chapman & Hall, London, 1975).
- Ohtani, E. & Ringwood, A. *Earth planet. Sci. Lett.* **71**, 85-93 (1984).
- Ohtani, E., Ringwood, A. & Hibberson, W. *Earth planet. Sci. Lett.* **71**, 94-103 (1984).
- Maale, S. & Aoki, K. *Contr. Miner. Petrol.* **63**, 161-173 (1977).
- Basaltic Volcanism Study Project *Basaltic Volcanism on the Terrestrial Planets* (Pergamon, New York, 1981).
- Bass, J. D. & Anderson, D. L. *Geophys. Res. Lett.* **11**, 237-240 (1984).
- Bass, J. D. & Anderson, D. L. (in preparation).
- Stolper, E. M. *Contr. Miner. Petrol.* **74**, 13-27 (1980).
- Rigden, S., Ahrens, T. & Stolper, E. *Science* **226**, 1071-1074 (1984).
- Nisbet, E. & Walker, D. *Earth planet. Sci. Lett.* **60**, 105 (1982).
- Ringwood, A. E. in *Advances in Earth Sciences* (ed. Hurley, P.) 287-356 (MIT Press, Cambridge, 1966).
- Anderson, D. L. *Science* **213**, 82-89 (1981).
- Anderson, D. L. *Adv. Earth planet. Sci.* **12**, 301-318 (1982).
- Anderson, D. L. *Earth planet. Sci. Lett.* **57**, 13-24 (1982).
- Anderson, D. L. *Earth planet. Sci. Lett.* **57**, 1-12 (1982).
- Morgan, J. W. & Anders, E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6973 (1980).
- Ringwood, A. E. *Origin of the Earth and Moon* (Springer, New York, 1979).
- DePaolo, D. J. & Wasserburg, G. J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3056 (1979).
- Jacobsen, S. B. & Wasserburg, G. J. *J. geophys. Res.* **84**, 7411-7427 (1979).
- Menzies, M. & Murthy, V. R. *Nature* **283**, 634 (1980).
- Frey, F. & Prinz, M. *Earth planet. Sci. Lett.* **38**, 129-176 (1978).
- McCulloch, M. T., Jaques, A., Nelson, D. & Lewis, J. *Nature* **302**, 400 (1983).
- Anderson, D. L. in *Proc. 3rd int. Kimberlite Conf.* (ed. Kornprobst, J.) 395-403 (Elsevier, New York, 1984).
- Nataf, H.-C., Nakanishi, I. & Anderson, D. L. *Geophys. Res.* **11**, 109-112 (1984).
- Anderson, D. L. & Dziewonski, A. *Scient. Am.* **251** No. 4, 60-68 (1984).
- Walck, M. *Geophys. J. R. astr. Soc.* **76**, 697-723 (1984).
- Given, J. W. & Helmlinger, D. V. *J. geophys. Res.* **85**, 7183-7194 (1980).
- Anderson, D. L. *Science* **223**, 347-355 (1984).
- Regan, J. & Anderson, D. L. *Phys. Earth planet. Inter.* **35**, 227-263 (1984).
- Grand, S. & Helmlinger, D. V. *J. geophys. Res.* **89**, 11,465-11,475 (1984).
- Anderson, D. L. & Bass, J. D. *Geophys. Res. Lett.* **11**, 637-640 (1984).
- Sawamoto, H., Weidner, D., Sasaki, S. & Kumazawa, M. *J. geophys. Res.* **89**, 7852-7860 (1984).
- Sawamoto, H., Weidner, D. J., Sasaki, S. & Kumazawa, M. *Science*, **224**, 749-751 (1984).
- Akimoto, S. *Tectonophysics* **13**, 161-187 (1972).

50. Jeanloz, R. & Thompson, A. B. *Rev. Geophys. Space Phys.* **21**, 51-74 (1983).
51. Ringwood, A. E. & Major, A. *Earth planet. Sci. Lett.* **12**, 411-418 (1971).
52. Akaogi, M. & Akimoto, S. *Phys. Earth planet. Inter.* **15**, 90-106 (1977).
53. Akaogi, M. & Akimoto, S. *Phys. Earth planet. Inter.* **19**, 31-51 (1979).
54. Liu, L. in *The Earth: Its Origin, Structure and Evolution* (ed. McElhinny, M. W.) 177-202 (Academic, New York, 1979).
55. Anderson, D. L. & Toksöz, M. N. *J. geophys. Res.* **68**, 3483-3500 (1963).
56. Niazi, M. & Anderson, D. L. *Bull. seismol. Soc. Am.* **58**, 339-366 (1968).
58. Johnson, L. R. *J. geophys. Res.* **72**, 6309-6325 (1967).
59. Archambeau, C. B., Flinn, E. A. & Lambert, D. G. *J. geophys. Res.* **74**, 5825-5866 (1969).
60. Kanamori, H. *Bull. Earthq. Res. Inst. Tokyo Univ.* **45**, 657-678 (1967).
61. Jeanloz, R. *J. geophys. Res.* **86**, 6171-6179 (1981).
62. Lees, A. C., Bukowinski, M. S. T. & Jeanloz, R. *J. geophys. Res.* **88**, 8145 (1983).
63. Anderson, D. L., Sammis, C. & Jordan, T. *Nature of the Solid Earth* (ed. Robertson, E.) 41-66 (McGraw-Hill, New York, 1972).
64. Sammis, C., Anderson, D. L. & Jordan, T. *J. geophys. Res.* **75**, 4478-4480 (1970).
65. Weidner, D. *Geophys. Res. Lett.* **12**, 417-420 (1985).
66. Ringwood, A. *J. Geol.* **90**, 611-643 (1982).
67. Bock, G. & Ha, J. *Geophys. J. R. astr. Soc.* **77**, 593-615 (1984).
68. Whitcomb, J. H. & Anderson, D. L. *J. geophys. Res.* **75**, 5713-5728 (1970).
69. Sobel, P. thesis, Univ. Minnesota (1978).
70. Richter, F. M. *J. geophys. Res.* **84**, 6783-6795 (1979).
71. Giardini, D. & Woodhouse, J. H. *Nature* **307**, 505-509 (1984).
72. Isacks, B. & Molnar, P. *Rev. Geophys. Space Phys.* **9**, 103-174 (1971).
73. Hager, B. H. & Richards, M. A. *J. geophys. Res.* **89**, 5987-6002 (1984).
74. Liebermann, R. C. & Ringwood, A. E. *J. geophys. Res.* **78**, 6926-6932 (1973).
75. Adams, R. D. *Bull. seismol. Soc. Am.* **61**, 1441-1451 (1971).
76. Richards, P. G. Z. *Geophys.* **38**, 517-527 (1972).
77. Creager, K. & Jordan, T. *J. geophys. Res.* **89**, 3031-3050 (1984).
78. Christensen, V. & Yuen, D. *J. geophys. Res.* **89**, 4389-4402 (1984).
79. Kato, T. & Kumazawa, M. *Nature* **316**, 803-805 (1985).
80. Weidner, D. *Phys. Earth planet. Inter.* **40**, 65-70 (1985).
81. Hirahara, K. Ishikawa, Y. *J. Phys. Earth* **32**, 197-218 (1984).
82. Ringwood, A. E. *Origin of the Earth and Moon* (Springer, New York, 1979).
83. Ito, E. & Matsui, Y. *Earth planet. Sci. Lett.* **38**, 442-450 (1978).
84. Liu, L. *Geophys. Res. Lett.* **1**, 277-280 (1974).
85. Mao, H., Yagi, T. & Bell, P. *Yb. Carnegie Instn Wash.* **76**, 502-504 (1977).
86. Ito, E., Takahashi, E. & Matsui, Y. *Earth planet. Sci. Lett.* **67**, 238-248 (1984).
87. Jackson, I. *Earth planet. Sci. Lett.* **62**, 91-103 (1983).
88. Knittle, E. & Jeanloz, R. *Nature* (in the press).
89. Ringwood, A. *J. geophys. Res.* **67**, 4005-4010 (1962).
90. Reid, A. & Ringwood, A. *J. geophys. Res.* **80**, 3363-3370 (1975).
91. Ahrens, T. J., Anderson, D. L. & Ringwood, A. E. *Rev. Geophys.* **7**, 667-707 (1969).
92. Davies, G. F. & Anderson, D. L. *J. geophys. Res.* **76**, 2617-2627 (1971).
93. Liu, L. *Phys. Earth planet. Inter.* **11**, 289-298 (1976).
94. Yagi, T., Mao, H. K. & Bell, P. M. *Phys. Chem. Miner.* **3**, 97-110 (1978).
95. Dzierwowski, A. M. & Anderson, D. L. *Phys. Earth planet. Inter.* **25**, 297-356 (1981).
96. Woodhouse, J. H. & Dzierwowski, A. M. *J. geophys. Res.* **89**, 5853 (1984).
97. Clayton, R. & Comer, R. *EOS* **64**, 776 (1983).
98. Anders, E. & Ebihara, M. *Geochim. cosmochim. Acta* **46**, 2363-2380 (1982).
99. Weidner, D., Bass, J., Ringwood, A. & Sinclair, W. *J. geophys. Res.* **87**, 4740-4746 (1982).
100. Yagi, Y., Akaogi, M., Shimomura, O., Tami, H. & Akimoto, S., *U.S. Japan Seminar, High-Pressure Research in Geophysics and Geochemistry, Program with Abstracts, January 13-16, 1986*, 62-63 (1986).

ARTICLES

Spatial restriction in expression of a mouse homoeo box locus within the central nervous system

Alexander Awgulewitsch*, Manuel F. Utset†, Charles P. Hart*, William McGinnis‡ & Frank H. Ruddle**

Departments of * Biology, † Human Genetics and ‡ Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511, USA

A common feature of Drosophila homoeo box genes appears to be their spatially restricted expression patterns during morphogenesis. Using Northern blot analysis and in situ hybridization to mouse tissue sections, the spatially restricted expression of a newly identified mouse homoeo box locus, Hox-3, within the central nervous system of newborn and adult mice has been demonstrated.

THE morphogenetic events which result in the segmented body plan of the fruit fly *Drosophila melanogaster* have been subjected to detailed genetic analyses. The genes controlling the segmentation processes and the identity of different segments are referred to as segmentation and homoeotic genes, respectively¹⁻³. The recent cloning of some of these genes located in the bithorax (BX-C) and *Antennapedia* (ANT-C) gene complexes⁴⁻⁶ has allowed the study of their molecular structure. One outcome of these structural analyses has been the detection of cross-homology between different homoeotic and segmentation genes. The best characterized cross-homology is known as the homoeo box, which encodes a highly conserved protein domain⁷⁻¹¹ that appears to be preferentially localized in the *Drosophila* genome within loci controlling fly morphogenesis. How this conserved sequence is related to the homoeotic or segmentation functions of these morphogenetic loci is not yet known.

Highly conserved copies of homoeo box sequences are found in a wide variety of animals, including mammals^{7,12-14}. The possibility that the mammalian and fly genes might perform similar morphogenetic functions is so far supported only by superficial similarities in their time of expression during embryogenesis¹⁵⁻²¹, their tissue-specific expression during embryonic²⁰, postnatal²¹ and adult stages¹⁹⁻²¹, their transcriptional activity in differentiated embryonic carcinoma cells from mouse and man^{16,17,19}, as well as in the clustered organization of homoeo box genes in both *Drosophila* and mammals^{15,17,19,21}. In *D. melanogaster*, the transcripts encoded by different homoeo box genes accumulate in regionally localized subsets of cells during

embryonic and larval development^{10,22-27}. The detection of similar patterns of mouse homoeo box gene expression during embryogenesis would be expected should the genes perform functions analogous to those of the *Drosophila* homoeo box gene family.

We report here the structural analysis of a murine homoeo box which we have mapped to chromosome 15. Analysis of the transcriptional activity of this newly identified homoeo box locus provides evidence for spatially restricted expression in a specific anterior-posterior region of the mouse central nervous system.

Isolation and structure of clone

A library of mouse genomic DNA in λ Charon 28 was screened under reduced stringency conditions with homoeo box-containing probes derived from the *Antennapedia* (*Antp*) and *Ultrabithorax* (*Ubx*) genes of *Drosophila*⁷. Nine separate recombinants were isolated after screening 10^6 λ clones (approximately three genome equivalents). The results presented here concern one of these clones, originally named λ Mo-EA. A restriction map of this clone obtained with six different restriction endonucleases is shown in Fig. 1a. The 2.2-kilobase (kb) *Bam*HI-*Eco*RI restriction fragment at the left end of the phage insert contains the homoeo box homology and was subcloned in the pSP64 vector system. The subclone, designated pMo-EA, was analysed for further restriction sites with two additional restriction enzymes, *Hae*III and *Hinf*I (Fig. 1a). The homology to the homoeo box sequences from *Drosophila* (*Ubx* and *Antp*) is confined to a 320-base-pair (bp) *Hae*III-*Hae*III fragment in

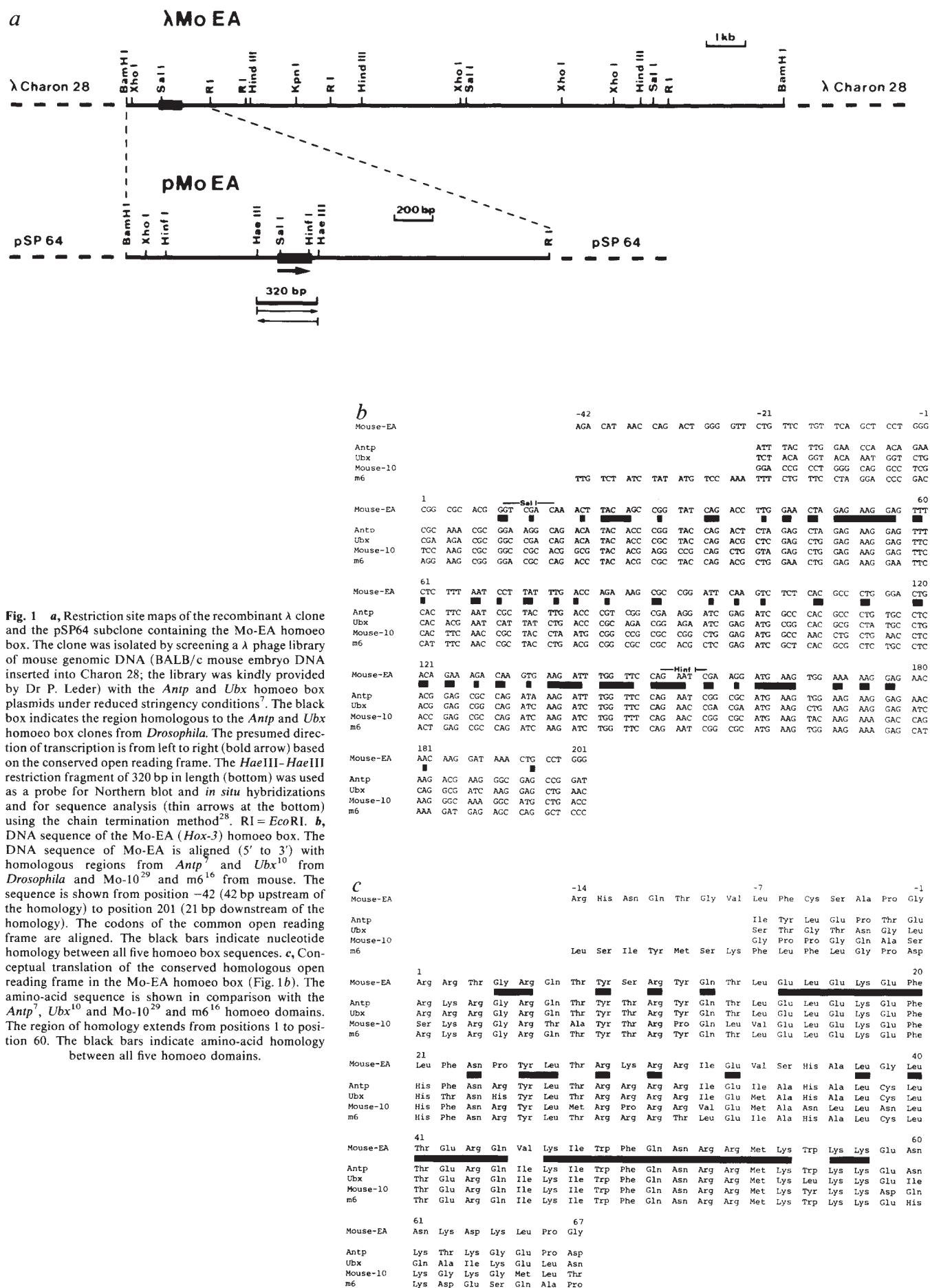


Fig. 1 **a**, Restriction site maps of the recombinant λ clone and the pSP64 subclone containing the Mo-EA homoeo box. The clone was isolated by screening a λ phage library of mouse genomic DNA (BALB/c mouse embryo DNA inserted into Charon 28; the library was kindly provided by Dr P. Leder) with the *Antp* and *Ubx* homoeo box plasmids under reduced stringency conditions⁷. The black box indicates the region homologous to the *Antp* and *Ubx* homoeo box clones from *Drosophila*. The presumed direction of transcription is from left to right (bold arrow) based on the conserved open reading frame. The *Hae*III-*Hae*III restriction fragment of 320 bp in length (bottom) was used as a probe for Northern blot and *in situ* hybridizations and for sequence analysis (thin arrows at the bottom) using the chain termination method²⁸. RI = *Eco*RI. **b**, DNA sequence of the Mo-EA (*Hox-3*) homoeo box. The DNA sequence of Mo-EA is aligned (5' to 3') with homologous regions from *Antp*⁷ and *Ubx*¹⁰ from *Drosophila* and Mo-10²⁹ and m6¹⁶ from mouse. The sequence is shown from position -42 (42 bp upstream of the homology) to position 201 (21 bp downstream of the homology). The codons of the common open reading frame are aligned. The black bars indicate nucleotide homology between all five homoeo box sequences. **c**, Conceptual translation of the conserved homologous open reading frame in the Mo-EA homoeo box (Fig. 1b). The amino-acid sequence is shown in comparison with the *Antp*⁷, *Ubx*¹⁰ and Mo-10²⁹ and m6¹⁶ homoeo domains. The region of homology extends from positions 1 to position 60. The black bars indicate amino-acid homology between all five homoeo domains.

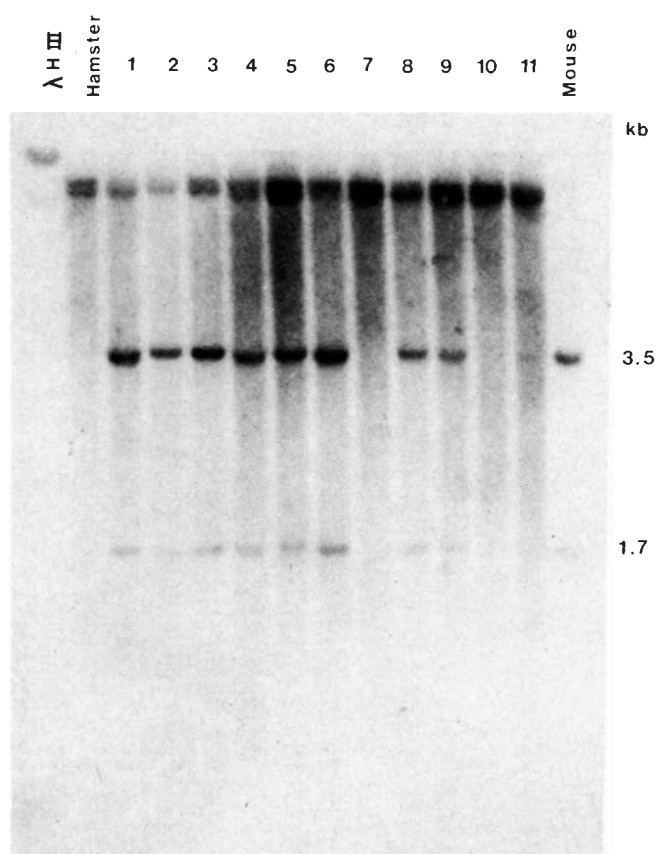


Fig. 2 Southern blot hybridization of mouse, hamster and mouse $\times$ hamster hybrid cell line DNAs with the *Hox-3* DNA probe, pMo-EA. Approximately 15 μ g of *Pst*I-digested DNA was loaded per track. The probe was nick translated in the presence of 32 P-dCTP to a specific activity of 10^8 c.p.m. μ g $^{-1}$ and hybridized as reported previously²⁹. Bacteriophage λ DNA digested with *Hind*III (λ HIII) was used for relative molecular mass determinations and the sizes of the mouse-specific hybridizing fragments are shown on the right. Hybrids listed from 1-11 are 2-1, 2-1-1, 2-1-7, 12G7, 4B31, 2a2c2, R44, EcM4, 10bs2, Mae32 and CEC, respectively. They are described in McGinnis *et al.*²⁹ and Pratcheva *et al.*³⁹.

the middle of the pMo-EA insert. This fragment was cloned in M13 mp11 in both orientations and the nucleotide sequence was determined using the dideoxy chain termination method²⁸. The 180-bp DNA sequence of the mouse EA homoeo box with 42-bp flanking sequence upstream and 21-bp sequence downstream is shown in Fig. 1b. Identification of the *Sall* and *Hinf*I restriction endonuclease sites close to the 5' end (*Sall*, base positions 11-16) and near the 3' region (*Hinf*I, base positions 150-154) within the homoeo box allows the 5' $\rightarrow$ 3' orientation on the restriction map (Fig. 1a). Comparison with the *Antp* and *Ubx* homoeo box DNA sequences from *Drosophila*^{7,10} shows ~73% and ~68% homology, respectively, while the homology to mouse 10 (ref. 29) and m6 from mouse¹⁶ is considerably lower, ~56% and ~66%, respectively.

The conceptual translation of the mouse EA homoeo box and its comparison with *Antp*⁷, *Ubx*¹⁰, mouse-10 (Mo-10)²⁹ and m6¹⁶ are shown in Fig. 1c. The homology of the amino-acid sequence of mouse EA to each of the four other homoeo domains is 83% to *Antp*, 80% to *Ubx*, 62% to Mo-10 and 78% to m6. The striking conservation in the amino-acid sequence is restricted to the homoeo domain, encoded by nucleotides 1-180. Out of 60 amino acids, 35 are identical at all five loci from mouse and fly.

Chromosomal mapping

Twenty-five mouse $\times$ hamster hybrid cell lines were screened for the presence or absence of mouse sequences cloned in pMo-EA. These cell lines contain the entire complement of hamster chromosomes and differing subsets of mouse chromosomes. The presence of a particular mouse chromosome is determined by

a combination of isozyme, DNA marker and karyotypic analyses.

Southern blot hybridizations using pMo-EA as a probe and mouse, hamster and mouse $\times$ hamster hybrid cell line DNAs digested with a variety of restriction endonucleases identified *Pst*I as an enzyme that distinguished the mouse EA hybridizing fragments from the cross-hybridizing hamster fragments. pMo-EA hybridizes with 3.5-kb and 1.7-kb *Pst*I fragments in a mouse genomic Southern blot. Cross-hybridizing bands of approximately 10 kb and 12 kb are present in the hamster genome using the mouse probe. A representative Southern blot is shown in Fig. 2. Scoring the mouse $\times$ hamster hybrid cell lines for the presence of the mouse-specific fragments assigns this genomic locus to mouse chromosome 15. A summary of the hybrid cell line blot hybridization results is shown in Table 1. The chromosomal location of mouse EA is different from the position of

Table 1 Southern blot hybridization results for mouse $\times$ hamster hybrid cell line DNAs probed with *Hox-3* probe pMo-EA

Mouse chromosome	Segregation of Mo-EA in mouse-hamster hybrids (presence of chromosome/blot hybridization)					
	Concordant			Discordant		
	+/+	-/-	Total	+/+	-/+	Total
1	12	6	18	1	6	7
2	13	5	18	2	3	5
3	11	7	18	0	4	4
4	13	5	18	2	5	7
5	6	6	12	1	12	13
6	13	6	19	1	5	6
7	16	3	19	4	2	6
8	12	7	19	0	6	6
9	10	4	14	3	7	10
10	14	6	20	0	5	5
11	0	7	7	0	18	18
12	15	3	18	4	2	6
13	4	3	7	0	2	2
14	5	6	11	1	13	14
15	6	3	9	0	0	0
16	3	2	5	1	3	4
17	17	4	21	3	3	6
18	1	6	7	1	10	11
19	8	5	13	2	6	8
X	8	2	10	5	7	12

The murine chromosomal complement of the 25 hybrid cell lines was determined by karyotyping and/or the analysis of isozyme and DNA markers. The presence or absence of mouse chromosome 15 is concordant with the presence or absence of the *Hox-3* diagnostic mouse bands.

two other homoeo box loci identified previously, *Hox-1* on chromosome 6²⁹, and *Hox-2* on chromosome 11^{30,31}. Keeping the same scheme of nomenclature, we call the newly identified mouse EA locus on chromosome 15, *Hox-3*.

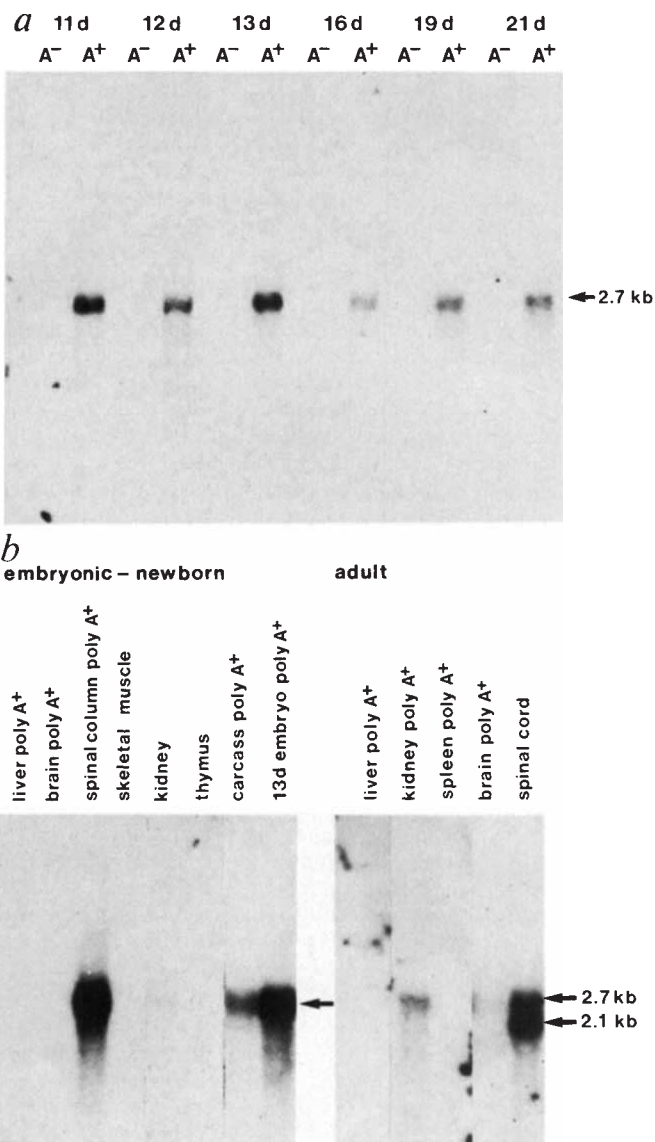
Hox-3 encodes tissue-specific transcripts

The expression of homoeo box genes in *D. melanogaster* follows a spatially and temporally restricted pattern. This has been most effectively demonstrated by *in situ* hybridizations to serial sections of embryos and larvae using genomic or complementary DNA probes^{10,22-27}. At mid-to-late stages of embryogenesis, different homoeotic transcripts are expressed in distinct regions of the fly central nervous system.

We have used both Northern blot analysis and *in situ* hybridization to histological sections to test transcript distributions in different tissues and in different anterior/posterior positions in the mouse body. In the first experiment, transcription from the *Hox-3* locus was tested at different stages of embryonic development. Poly(A)⁻ and poly(A)⁺ RNAs were isolated from total embryos at 11, 12, 13, 16, 19 and 21 days post coitum and hybridized with the nick-translated pMo-EA probe from *Hox-3*.

Fig. 3 a, Northern blot hybridizations of *Hox-3* to poly(A)⁻ and poly(A)⁺ RNAs (shown as A⁻ and A⁺) isolated from whole mouse embryos of different gestation stages, starting with day 11 (11d). At gestation stage day 21 (21d) the mice were born and the newborn mice were used for the experiment. **b**, Northern blot hybridizations of *Hox-3* to poly(A)⁺ and total RNAs isolated from liver, brain, spinal column, skeletal muscle, kidney, thymus and carcass dissected from newborn CD-1 mice (left), and to poly(A)⁺ RNAs from liver, kidney, spleen, brain and spinal cord (total RNA) from adult male and female CD-1 mice (right). Poly(A)⁺ RNA from whole embryos of gestation stage day 13 (last track on the left) served as a positive control.

Methods. The breeding of mice (CD-1), the dissection of embryos and tissues and the isolation of RNA using the guanidinium thiocyanate CsCl method⁴⁰ were done essentially as described previously¹⁵. The spinal column from newborn mice was dissected by cutting it directly behind the skull and by separating it from limbs and ribs on either side. Most of the muscle and connective tissue was removed before the RNA was isolated. For the isolation of RNA from the spinal cord of adult animals, the spinal column was first dissected in the same way. The spinal cord was separated from the vertebrae by nicking the spinal column with 5–10 transverse cuts set along its anterior-posterior axis, and by pulling the vertebrae off the cylindrical nerve cord in the longitudinal direction. By this procedure it was possible to prepare the entire nerve cord from the atlas region to the region of the sacral vertebrae. Carcass (b) means the whole body of a newborn mouse after removal of the spinal column. Eight µg of each poly(A)⁻ and poly(A)⁺ RNA sample (a) and 8 µg of poly(A)⁺ RNA or 20 µg of total RNA (b) were separated in formaldehyde gels and transferred to nitrocellulose gels according to Maniatis *et al.*⁴¹. The RNA filters were prehybridized for 8–15 h at 42 °C in 50% formamide, 5 × SSC, 50 mM sodium phosphate (pH 6.6), 0.1% SDS, 5 × Denhardt's solution⁴² and 100–200 µg ml⁻¹ of denatured sonicated salmon sperm DNA (Sigma). The hybridization with nick-translated plasmid DNA (pMo-EA) was carried out under the same conditions with 10% dextran sulphate added for 20–24 h. The concentration of denatured ³²P-labelled probe DNA (specific activities ~2 × 10⁸ c.p.m. per µg DNA) was about 3–5 ng ml⁻¹. The filters were washed in 3 × SSC, 0.1% SDS at room temperature for 2 × 10 min and in 0.2 × SSC, 0.1% SDS at 65 °C for 3 × 20 min. The filters were autoradiographed with XAR-5 film from Kodak and an intensifying screen for 2 days at -70 °C. As a control (b) the filters were stripped with boiling 0.1 × SSPE, 0.1% SDS and hybridized with a ³²P-labelled cDNA clone of the human hypoxanthine phosphoribosyl transferase (*hprt*) gene⁴³. Each RNA sample displayed a single band at the same position, but of different intensity, indicating the integrity of the RNAs.



The results show an abundant transcript of approximately 2.7 kb in size, present only in the poly(A)⁺ RNA fractions of all these embryonic stages (Fig. 3a). No hybridization signals were detectable with the poly(A)⁻ RNAs even after longer autoradiographic exposures. Embryonic stages earlier than day 11 post coitum have not yet been tested.

In order to investigate the tissue distribution of the *Hox-3* transcript(s), we prepared RNA blots using RNA isolated from a number of different tissues and organs from newborn and adult animals, and hybridized them with the pMo-EA probe (Fig. 3b). Analysis of the RNAs from newborn mice (Fig. 3b, left) showed that the strongest hybridization signal was obtained with the poly(A)⁺ RNA isolated from the spinal column, indicating that at this stage of development a major site of *Hox-3* transcript accumulation is in cells of the spinal column. The size of the hybridizing transcript of 2.7 kb is identical to the size of the abundant transcript in 13-day embryos. A less abundant transcript of the same size is present in the RNA samples from kidney and from the carcass remaining after removal of the spinal column. The weak signal with the kidney was obtained with total RNA, while the stronger signal with the carcass was obtained with poly(A)⁺ RNA. Since there was no hybridization signal detectable with the RNAs from the other tissues and organs tested, including liver, brain, skeletal muscle and thymus, we suspect that much of the hybridization signal in the carcass is due to kidney-specific expression of the *Hox-3* locus.

The distribution of *Hox-3* transcripts in RNA samples from adult tissues (Fig. 3b, right) is consistent with our data concerning *Hox-3* expression in newborn mice and allows us to draw additional conclusions. While liver, spleen, skeletal muscle (data not shown) and brain appear to be negative for *Hox-3* expression, the kidney and the spinal cord are positive. A strong signal is obtained with the unfractionated total RNA from the spinal cord, in contrast to the weaker hybridization signal obtained with the poly(A)⁺ RNA from the kidney, suggesting a much higher level of *Hox-3* expression in the adult spinal cord than in the adult kidney. The data allow us to assign a specific part of the central nervous system—the spinal cord—as at least one major site for *Hox-3* expression in adult mice. Although the tissue distribution of *Hox-3* transcripts appears to be the same in newborn and adult, in adult spinal cord an additional transcript of about 2.1 kb in size appears to be present in a higher copy number than the 2.7-kb transcript. This 2.1-kb transcript is not detectable in the other RNA samples.

Localized accumulation of *Hox-3* transcripts

If the *Hox-3* locus functions in a manner similar to *Drosophila* homoeotic genes, one would expect spatial restrictions of the transcripts that we have detected within the spinal cord. We have tested this prediction by two different experimental approaches: first by using Northern blot analysis of RNA isolated from three distinct regions of both the newborn spinal

column and the adult spinal cord, and second, by using *in situ* hybridization to tissue sections.

For the Northern blot hybridizations we used the 320-bp *Hae*III fragment which contains the entire *Hox-3* homoeo box and 140 bp of flanking sequences (Fig. 1*a, b*). In mRNA samples from the whole spinal column of newborn mice and from the spinal cord of adults (Fig. 4*c*, tracks Se+Sa, top) the 2.7-kb transcript was detected in both newborn and adult tissue, and the 2.1-kb RNA only in adult.

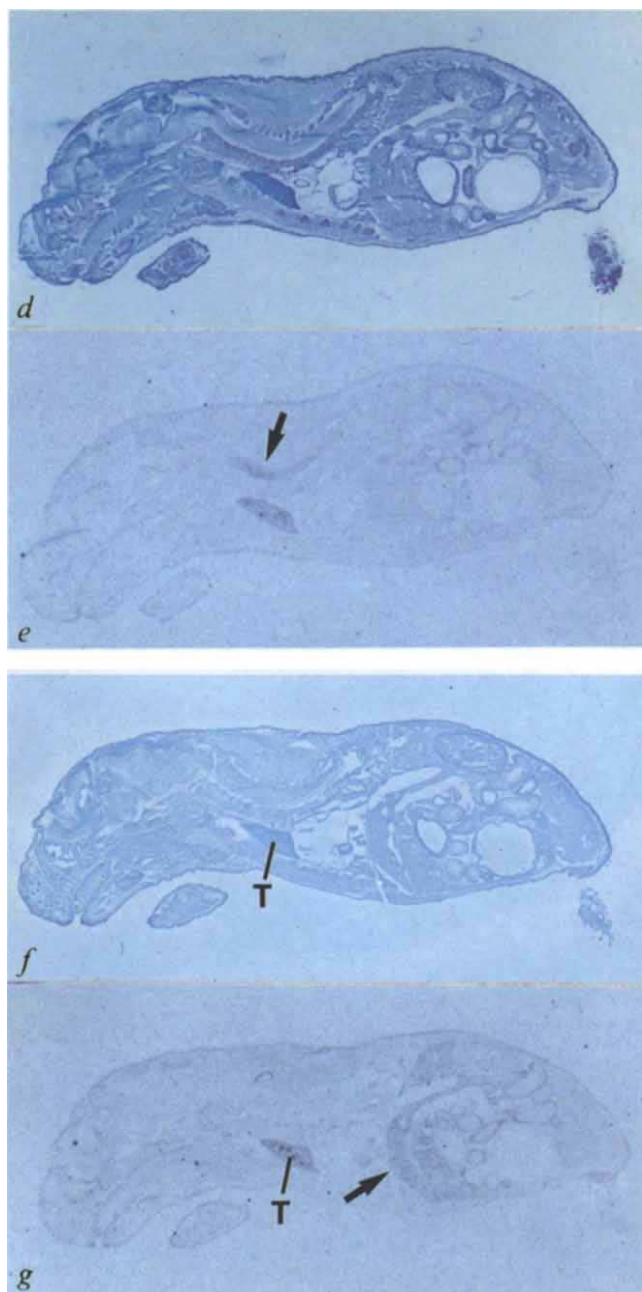
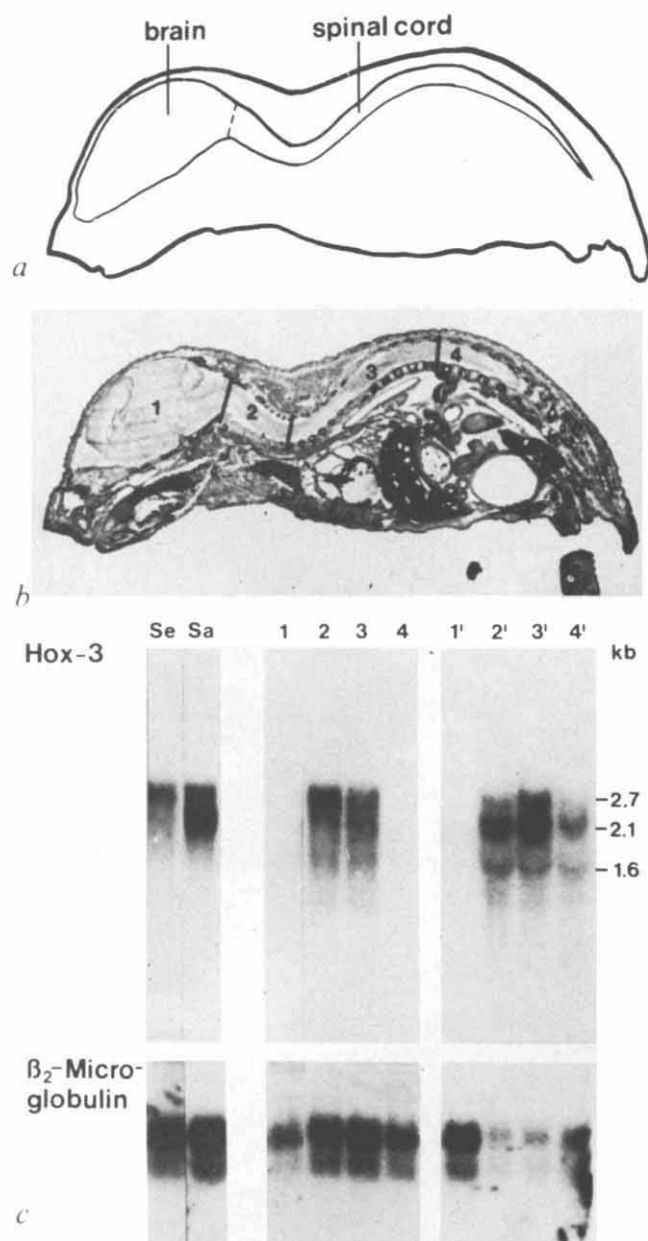
The results obtained with RNA samples from the newborn brain and from the cervical, the thoracic, and the lumbar/sacral/caudal regions of the newborn spinal column are shown in Fig. 4*c*, tracks 1–4. The photograph of a sagittal section of a newborn mouse (Fig. 4*b*) and a schematic drawing of this photograph (Fig. 4*a*) depict the regions of the central nervous system or spinal column from which the four different RNA samples were isolated (Fig. 4 legend). Strong hybridization signals were obtained with the RNA samples from the cervical and thoracic regions (tracks 2 and 3 in Fig. 4*c*, top), while the hybridization to the samples from the brain and lumbar region (tracks 1 and 4) appeared to be either negative (brain) or weak (lumbar). In the control experiment (Fig. 4*c*, bottom) carried out with the same filter and using a mouse β_2 -microglobulin probe^{32,33}, all three RNA samples from the newborn spinal column showed roughly equal signals, while the signal with the RNA from newborn brain was considerably weaker, though easily detectable. These data suggest a spatially restricted accumulation of *Hox-3* transcripts in the anterior regions of the spinal column from newborn mice. In addition to the 2.7-kb transcript, two other RNAs, 2.1 kb and 1.6 kb in size, were detected (Fig. 4*c*, tracks 2, 3). The 2.1-kb RNA presumably corresponds to the transcript also detected in the adult spinal cord (Figs 3*b*; 4*c*, track Sa, top), where it is the most abundant *Hox-3* transcript. The weak signal at the 1.6-kb position was not detected in the experiments described above (Figs 3*a, b*; 4*c*, tracks Se, Sa, top). Since we see a relatively stronger signal at the corresponding position by hybridizing the 320-bp *Hae*III fragment from *Hox-3* to total RNAs from adult spinal cord (Fig. 4*c*, tracks 2'–4', top), but not upon hybridization to the poly(A)⁺ RNA from adult spinal cord (Fig. 4*c*, track Sa, top), it is possible that these bands result from hybridization to heterogeneous RNA fragments which have been compressed into this region close to the 18S ribosomal RNA²⁶.

The spatial restriction of *Hox-3* transcripts in the central nervous system of adult mice was also tested by RNA blot analysis. A messenger RNA sample from the brain and samples of equal amounts of total RNA from the cervical, thoracic and lumbar regions of adult spinal cord were probed with the 320-bp *Hae*III fragment containing the *Hox-3* homoeo box (Fig. 4*c*, tracks 1'–4', top). No hybridization to the mRNA from the brain is detected, while all three RNA samples from the different spinal cord regions show hybridization to the 2.7- and 2.1-kb transcripts. High signal intensities are detected with the RNAs from the cervical and thoracic regions of the spinal cord, in contrast to the considerably weaker hybridization signals obtained with the RNA sample from the lumbar region. The significance of these signal intensity differences is supported by the control hybridization with the β_2 -microglobulin probe (Fig. 4*c*, tracks 1'–4', bottom), showing equal intensities in the tracks corresponding to the cervical, thoracic and lumbar regions, which indicates equal amounts of RNA in each of these tracks.

We have also examined the spatial restriction of *Hox-3* expression in newborn mice by hybridizing the 320-bp homoeo box probe to whole body serial sections (Fig. 4*d, e*). As a control we used a probe containing sequences from the mouse α -fetoprotein (AFP) gene, which is strongly expressed in newborn mouse liver³⁴ (Fig. 4*f, g*). Comparison of an autoradiograph of a sagittal section hybridized with the *Hox-3* probe (Fig. 4*e*) with a photograph of the stained section (Fig. 4*d*) indicates that

Fig. 4 (opposite) *a*, Schematic drawing of a sagittal section of a newborn mouse shown in *b*. *b*, Photograph of a Giemsa-stained sagittal section of a newborn mouse. For the isolation of RNAs from different regions of the central nervous system (CNS), the spinal column was separated from the brain and cut into three different sections as indicated by the black lines. The numbers label the separated regions of the CNS or the spinal column, respectively, from which RNAs have been isolated and refer directly to the labels used for the RNA blot hybridizations with the *Hox-3* probe in *c*; 1, brain; 2, cervical region; 3, thoracic region; 4, lumbar/sacral region. *c*, Top, Northern blot hybridization of *Hox-3* to RNA samples from the brain, and from different sections of the newborn spinal column and the adult spinal cord. The different tracks represent the following RNAs with the amounts (μ g) determined by A_{260} measurement; Se, poly(A)⁺ RNA from whole newborn spinal column (8 μ g); Sa, poly(A)⁺ RNA from whole adult spinal cord (8 μ g); 1–4, poly(A)⁺ RNAs from the newborn brain (1) and from the cervical (2), the thoracic (3) and the lumbar/sacral/caudal (4) region of the newborn spinal column (8 μ g each); 1', poly(A)⁺ RNA from the adult brain (8 μ g); 2'–4', total RNAs from the cervical (2'), thoracic (3') and lumbar/sacral (4') region of the adult spinal cord (40 μ g each). Bottom, control hybridization of the RNA blot represented on the top with a mouse β_2 -microglobulin probe (kindly provided by Dr P. Leder). *d*, Bright-field view of a sagittal section of a newborn mouse that was hybridized with the *Hox-3* probe. *e*, Autoradiograph obtained after exposure of the section in *d* to Kodak XAR-5 film for 20 days. The arrow indicates the hybridization signal resulting from *Hox-3* transcripts in the spinal cord. The apparent signal in the head region is due to a fold in the section at that position. The artefactual labelling of the thymus is discussed in the text. *f*, Bright-field view of a sagittal section of a newborn mouse that was hybridized with a mouse AFP gene probe (kindly provided by Dr S. Tilghman). *g*, thymus. *g*, Autoradiograph obtained after exposure of the section in *f* to Kodak XAR-5 film for 20 days. The arrow indicates the hybridization signal resulting from AFP transcripts in the liver. T, thymus.

Methods. Dissections of nervous tissues, RNA isolations and Northern blot hybridizations: The dissection of the brain and spinal column or the preparation of the spinal cord in the case of adult animals was done as described in Fig. 3 legend. The three different sections of the cervical, thoracic and lumbar/sacral/caudal region of the spinal column were obtained by cutting behind the 7th and 20th vertebrae from anterior to posterior as indicated in *b*. The isolation of RNAs, electrophoresis, blotting to nitrocellulose and hybridizations were carried out as referred (Fig. 3). The *Hox-3* probe was the 320-bp *Hae*III fragment shown in Fig. 1*a*; the β_2 -microglobulin gene probe has been described previously^{32,33}. ***In situ* hybridization.** The *in situ* hybridization to frozen sections of newborn mice was performed by a modification of the technique of Lawrence and Singer⁴⁴. Briefly, newborn mice were killed with chloroform and quick-frozen in OCT embedding medium (Miles Laboratories) on a cryostat specimen holder by immersing the holder into liquid nitrogen. After warming to -20°C , the frozen blocks were sectioned at 8–10 μm on a SLEE cryostat. The sections were picked up onto acid-cleaned microscope slides subbed according to Gall and Pardue⁴⁵, except that 0.3% gelatin was used. The sections were then fixed and dehydrated according to the method of Hafen *et al.*⁴⁶. The plasmid pAFP-440, a pSP65 derivative, contains ~130 bp of AFP coding sequences within a 440-bp *Hinc*II fragment of mouse genomic DNA. pAFP-440 and the 320-bp *Hae*III fragment of *Hox-3* were labelled to specific activities approaching 1×10^8 d.p.m. with ^{35}S -labelled deoxynucleotide 5'-(α -thio)triphosphates (Amersham) as follows. Three parallel nick-translation reactions with increasing concentrations of DNase I (Sigma) were set up for each probe. In each reaction, 0.5 μg of DNA was suspended in 30 μl 1 $\times$ nick-translation buffer⁴¹ containing ~50 μCi each of ^{35}S -labelled dATP and dCTP, and 20 μM each of unlabelled dGTP and dTTP. DNase I concentrations used were 0.5 ng ml⁻¹, 1.0 ng ml⁻¹ and 2.0 ng ml⁻¹. 20 units of DNA polymerase I (Boehringer) were added and the reaction mixes were incubated at room temperature for 60 min. The reactions were stopped by the addition of 100 μl of 0.1 M NaCl, 10 mM Tris-HCl (pH 8.0), 2 mM EDTA and 0.02% SDS, and the probes were separated from the unincorporated nucleotides by passage through a Sephadex-G50 spin column⁴¹. The probes were then boiled for 1 min and stored at -20°C . Just before hybridization, the probes were suspended at a concentration of 1 ng μl^{-1} in freshly prepared hybridization buffer containing 0.5 ng ml⁻¹ *Escherichia coli* transfer RNA, 1 $\times$ Denhardt's solution⁴², 50% deionized formamide, 0.75 M NaCl, 10 mM Tris-HCl (pH 7.5), 2 mM EDTA, 10% dextran sulphate, and 10 mM dithiothreitol. 50 μl of hybridization solution was applied to each newborn mouse section and covered with a 24 $\times$ 40-mm coverslip. The coverslips were sealed with rubber cement, and then the slides were heated to 80 $^\circ\text{C}$ for 2 min to denature the probe⁴⁷. Hybridization was carried out in a moist chamber at 37 $^\circ\text{C}$ for 4–5 h. The sections were washed by first removing the rubber cement and then allowing the coverslips to float off in a solution of 2 $\times$ SSC, 50% formamide, and 0.1% 2-mercaptoethanol. The sections were then incubated for at least 30 min in the same buffer at 37 $^\circ\text{C}$, followed by at least 30 min in 1 $\times$ SSC, and 50% formamide at 37 $^\circ\text{C}$, and finally by at least 30 min in 1 $\times$ SSC at room temperature with gentle shaking. The sections were then dehydrated in 70% and 95% ethanol (2 $\times$ 5 min, each), air dried and exposed to Kodak XAR-5 film for 20 days. The sections were then immersed in Kodak NTB-2 emulsion as described by Hafen *et al.*⁴⁶, except that the emulsion was diluted 1:1 with H₂O, and autoradiography was allowed to proceed for 15 days at 4 $^\circ\text{C}$ in a dry, light-tight box. The slides were developed in Kodak D-19 developer for 2 min, rinsed briefly in H₂O, and fixed in Kodak rapid-fix for 4 min, all at about 15 $^\circ\text{C}$. The sections were stained with Giemsa and examined by light microscopy.



Hox-3 transcripts accumulate in cells of the spinal cord. In numerous serial sections that have been probed in this manner, *Hox-3* transcripts have been detected throughout the posterior region of the spinal cord, with an anterior limit of transcript accumulation observed at approximately the level of the third cervical vertebra. A close-up view of this region of the spinal cord is depicted in Fig. 6. Note the decrease in specific signal over cell bodies anterior to the C-3/C-4 region. No obvious morphological transition occurs at this expression boundary. The highest signal levels are always observed adjacent to the boundary in the cervical region and in the thoracic region of the spinal cord. No signals forward of this boundary in any part of the central nervous system have been observed at any section depth. The cells of the thymus are also strongly labelled with the *Hox-3* probe, but this signal is apparently artefactual because the AFP probe labelled the thymus as strongly as the *Hox-3* probe (Fig. 4f, g), and because Northern blots of thymus RNA indicate no accumulation of *Hox-3* transcripts (see Fig. 3b). We observed no detectable labelling of kidney cells with the *Hox-3*

probe, presumably due to the relative insensitivity of the *in situ* method versus Northern blot analysis.

To better characterize the localization of transcripts in the newborn spinal cord and to confirm the presumed direction of transcription, single-stranded probes were used for *in situ* hybridizations to cross-sections of three different regions of the newborn spinal column: just posterior to the second, sixth and twentieth vertebrae (Fig. 5). For these experiments, the homoeo box-containing 320-bp *Hae*III fragment cloned in the Gemini vector system was used as a template for the synthesis of 35 S-labelled RNA transcripts from both strands. A strong specific signal was obtained only from hybridizations to cross-sections of the spinal cord region just posterior to the sixth vertebra and only using the probe detecting transcripts synthesized in the direction predicted by the orientation of the homoeo box (Fig. 5b'). Comparing the autoradiograph with the stained section (Fig. 5b) indicates an accumulation of *Hox-3* transcripts in the centre of the spinal cord, whereas the spinal ganglia do not appear specifically labelled.

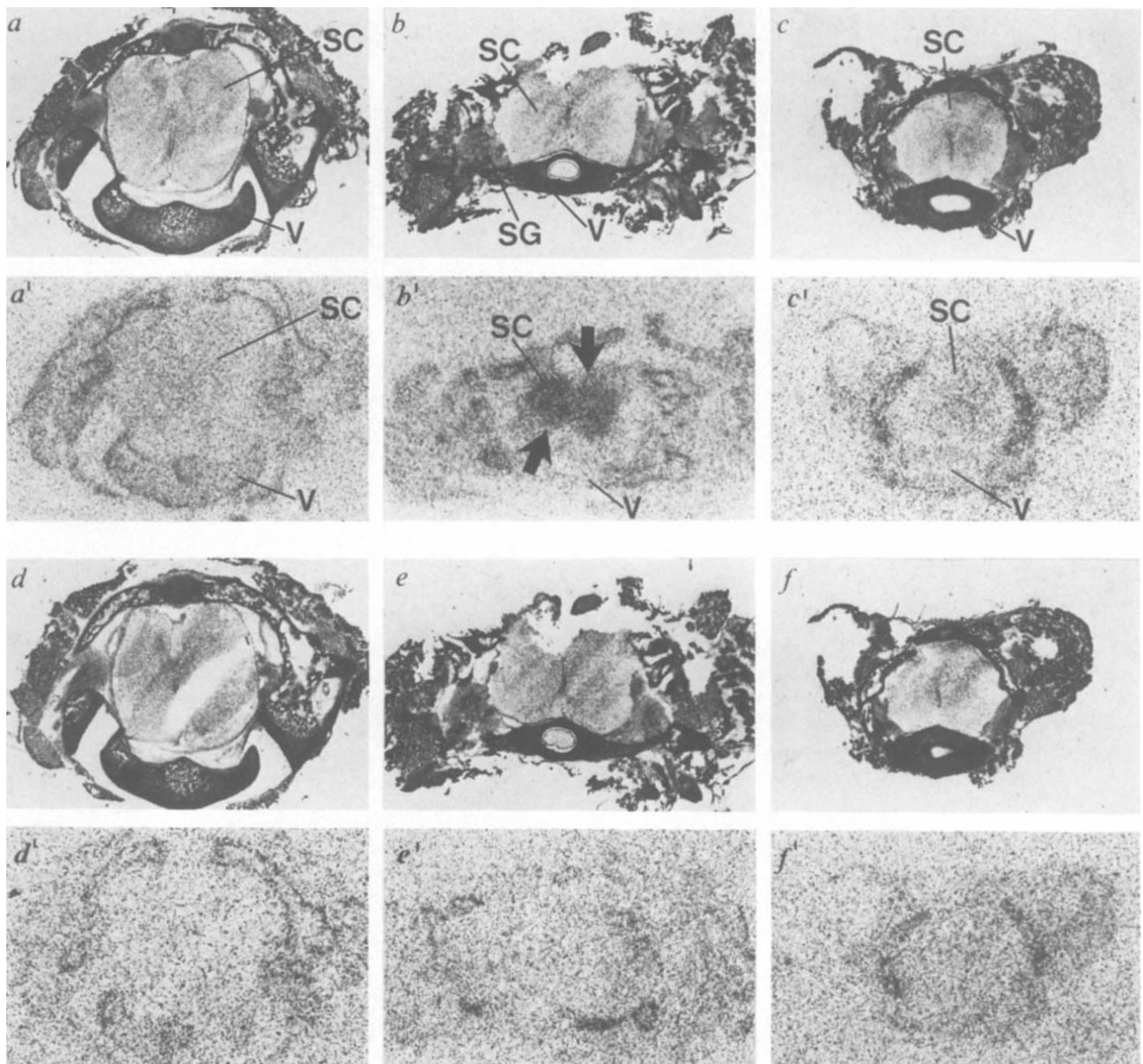


Fig. 5 *a-c*, Bright-field views of Giemsa-stained sections of the newborn spinal cord from regions just posterior to the 2nd (*a*), 6th (*b*) and 20th (*c*) vertebrae hybridized with ^{35}S -labelled RNA complementary to the coding strand of *Hox-3*. All three sections were hybridized on the same microscope slide. The dorsal side is at the top of each photograph. V, vertebral body; SC, spinal cord; SG, spinal ganglion. *a'-c'*, Autoradiographs obtained after exposure of the sections shown in *a, b, c* to Kodak RP-5 film for 48 h. The arrows in *b'* (spinal cord region just posterior to the 6th vertebra) point to the specific signal in the centre of the spinal cord. The surrounding nonspecific labelling is mainly concentrated over dense cartilaginous tissue. *d-f*, Bright-field views of Giemsa-stained cross-sections of the newborn spinal cord from regions just posterior to the 2nd (*d*), 6th (*e*) and 20th (*f*) vertebrae, hybridized with ^{35}S -labelled RNA complementary to the anticoding strand of *Hox-3*. All three sections were hybridized on the same microscope slide and are from the same animal as those in *a, b* and *c*. *d'-f'*, Autoradiographs obtained after exposure of the sections in *a, b, c* to Kodak XRP-5 film for 48 h.

Methods. The 320-bp *Hae*III fragment (Fig. 1a) was cloned into the poly-linker region of pGem4 (Promega Biotec) by standard methods. High specific activity RNA transcripts ($\sim 10^8$ c.p.m. μg^{-1}) complementary to coding and anticoding strands were prepared essentially as described by Cox *et al.*⁴⁸, as modified by the manufacturer. Labelled nucleotide, ^{35}S -labelled UTP (400 Ci mol^{-1} , Amersham), was used at a concentration of 12 μM in the synthesis reaction. Hybridizations were performed for 6 h at 50 $^{\circ}\text{C}$ with a probe concentration of 0.25 $\mu\text{g ml}^{-1}$, essentially as described by Cox *et al.*⁴⁸. After allowing the coverslips to float off in a solution of 2 $\times$ SSC, 50% formamide, and 0.1% β -mercaptoethanol, wash conditions were as follows: 2 $\times$ SSC, 50% formamide, 0.1% β -mercaptoethanol, 50 $^{\circ}\text{C}$; 20 $\mu\text{g ml}^{-1}$ RNase, 0.5 M NaCl, 10 mM Tris-Cl (pH 8.0), 37 $^{\circ}\text{C}$; 2 $\times$ SSC, 50% formamide, 0.1% β -mercaptoethanol, 50 $^{\circ}\text{C}$; 1 $\times$ SSC, 50% formamide, 0.1% β -mercaptoethanol, 50 $^{\circ}\text{C}$; all for at least 30 min. Slides were then dehydrated and exposed to X-ray film.

Conclusions

If mouse homoeo box genes direct specific morphogenetic functions during embryonic development, one would predict spatial localization of their expression to specific body regions of the developing mouse. Studies from several laboratories have already demonstrated differential expression of individual homoeo box genes in different mouse tissues at several developmental stages^{19,20,21}. For example, *Hox-2.1* transcripts are abundant in the embryonic, newborn and adult central nervous system^{20,21}. As shown in this study, *Hox-3* transcripts also accumulate in the newborn and adult central nervous system,

and more importantly, *Hox-3* transcript abundance displays a striking spatially restricted localization along the longitudinal axis of the central nervous system (Figs 4, 5). The highest level of *Hox-3* transcript accumulation occurs in the posterior cervical and anterior thoracic region of the spinal cord. We detect no expression more anteriorly and only weak expression in more posterior regions of the central nervous system.

One of the distinguishing characteristics of many of the homoeo box-containing homoeotic genes of *Drosophila* is their localized expression within discrete regions of the ventral nerve cord of the developing fly^{23,27}. For the *Antp* gene, this restricted pattern of expression persists at least until the third instar larval

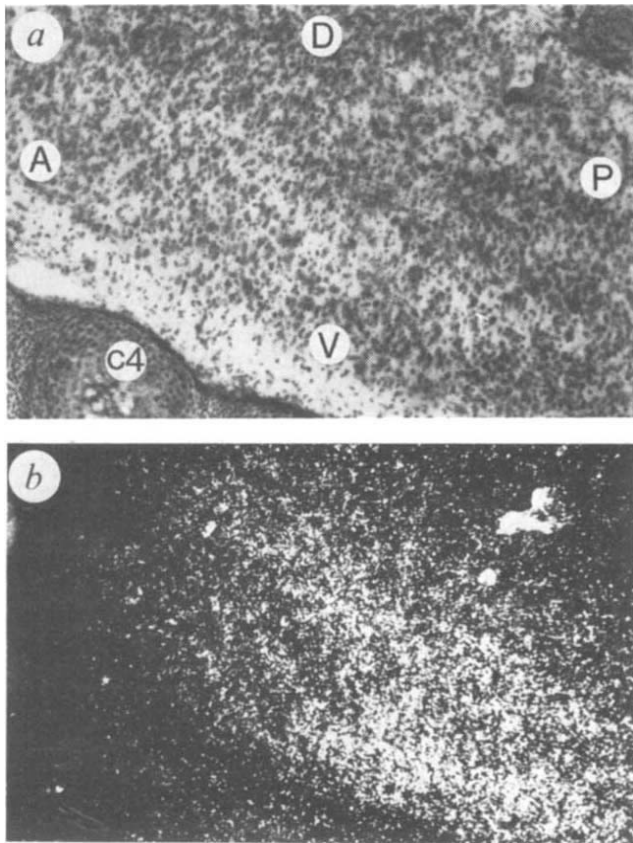


Fig. 6 Bright-field (a) and dark-field (b) views of the anterior-most region of specific labelling of the spinal cord of the section depicted in Fig. 4d. A, anterior; D, dorsal; P, posterior; V, ventral; C4, fourth cervical vertebra. ($\times 69.3$).

stage³⁵. By Northern blot analysis *Antp* products can still be detected in late pupal stages and also in the adult fly⁶.

Although attempts to draw functional parallels between the pattern of *Hox-3* expression in newborn and adult mice and the patterns of expression of homoeotic genes during fruit fly embryogenesis are still admittedly speculative, the similarity of these patterns is dramatic. In addition, preliminary observations indicate that the observed localization of *Hox-3* expression occurs at least as early as the 13th day of embryogenesis (M.F.U. *et al.*, unpublished data). We therefore conclude that the observed patterns of expression of *Hox-3* and *Drosophila* homoeo box sequences within the respective central nervous systems of mice and flies is consistent with a developmental role for *Hox-3* within a specific region of the mouse body plan.

Morphogenetic loci might also be expected to show specific temporal patterns of transcriptional activity during development. *Hox-3* is transcriptionally active at all developmental stages tested so far, spanning the time between the 11th day of gestation and the adult mouse. With the *Hox-3* probe pMo-EA, we detected an abundant transcript of 2.7 kb at all of these stages (Fig. 3a, b). Although no obvious temporal differences in *Hox-3* expression were detected during gestation, a difference in *Hox-3* expression between newborn and adult mice was detected. This difference is indicated by a shift in the relative abundance of the 2.1-kb and 2.7-kb transcripts in RNA samples from newborn spinal column versus adult spinal cord (Figs 3b, 4c). We also detected relatively low levels of the 2.7-kb *Hox-3* transcript in RNA samples from newborn and adult kidney (Fig. 3b).

A number of previously identified genes controlling morphogenesis in *Drosophila*, such as *Deformed*, *infra-abdominal 2*, *infra-abdominal 7* (refs 11, 36) and *engrailed*¹⁰, have been isolated by virtue of their cross-homology to homoeo box probes. Attempting to correlate homoeo box genes with known morphogenetic loci in the mouse is much more difficult because of the relatively small number of murine mutations known to disrupt morphogenesis, and the low resolution of cytogenetic mapping. However, it may eventually be possible to reveal the suspected morphogenetic functions of murine homoeo box genes by analysing their genetic linkage to the known developmental mutations in mouse. In contrast to the previously characterized mouse homoeo boxes which are located on chromosomes 6²⁹ and 11^{30,31}, the *Hox-3* locus maps to mouse chromosome 15. In light of the expression of *Hox-3* in the spinal cord, it is interesting to note that the semi-dominant mutation velvet coat (*Ve*), which in the homozygous state disrupts neural tube formation³⁷, also maps to this chromosome³⁸.

The results reported here are entirely consistent with a morphogenetic role for the *Hox-3* locus. However, because of the descriptive and comparative nature of our data, this view must remain conjectural. The next stage of analysis must involve direct experimentation to test this hypothesis.

We thank Nadine McGinnis and Lois Nichols for technical assistance, Marie Siniscalchi for typing the manuscript, Suzy Pafka for photographic assistance, and Drs Howard Lieberman and Abraham Fainsod for reviewing the manuscript. We also thank Drs Phil Leder and Shirley Tilghman for cloned DNA, Dr Dimitrina Pravtcheva for hybrid cell line DNAs, and Dr Michael Levine for advice on *in situ* hybridization methodologies. M.F.U. received financial support from the Life and Health Insurance Medical Research Fund and the Medical Scientist Training Program. C.P.H. was supported by an NIH predoctoral training grant. This work was supported by NIH grant GM09966 to F.H.R. and NSF grant DCB8501822 to W.M.

Received 6 November 1985; accepted 7 February 1986.

1. Nusslein-Volhard, C. & Wieschaus, E. *Nature* **287**, 795–801 (1980).
2. Garcia-Bellido, A. *Am. Zool.* **17**, 613–629 (1977).
3. Lewis, E. B. *Nature* **276**, 565–570 (1978).
4. Bender, W. *et al. Science* **221**, 23–29 (1983).
5. Garber, R. L., Kuroiwa, A. & Gehring, W. J. *EMBO J.* **2**, 2027–2036 (1983).
6. Scott, M. P. *et al. Cell* **35**, 763–776 (1983).
7. McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A. & Gehring, W. J. *Cell* **37**, 403–408 (1984).
8. Scott, M. P. & Weiner, A. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1095–1099 (1984).
9. Poole, S. J., Kauvar, L. M., Drees, B. & Kornberg, T. *Cell* **40**, 37–43 (1985).
10. Fjose, A., McGinnis, W. & Gehring, W. J. *Nature* **313**, 284–289 (1985).
11. Regulski, M. *et al. Cell* **43**, 71–80 (1985).
12. Carrasco, A. E., McGinnis, W., Gehring, W. J. & DeRobertis, E. M. *Cell* **37**, 409–414 (1984).
13. Muller, M. M., Carrasco, A. E. & DeRobertis, E. M. *Cell* **39**, 157–162 (1984).
14. McGinnis, W. *Cold Spring Harb. Symp. quant. Biol.* **50**, 263–270 (1985).
15. Hart, C. P., Angewilewitsch, A., Fainsod, A., McGinnis, W. & Ruddle, F. H. *Cell* **43**, 9–18 (1985).
16. Colberg-Poley, A. M., Voss, S. D., Chowdhury, K. & Gruss, P. *Nature* **314**, 713–718 (1985).
17. Hauser, C. A. *et al. Cell* **43**, 19–28 (1985).
18. Joyner, A. L., Kornberg, T., Coleman, K. G., Cox, D. R. & Martin, G. R. *Cell* **43**, 29–37 (1985).
19. Colberg-Poley, A. M. *et al. Cell* **43**, 39–45 (1985).
20. Jackson, I. J., Schofield, P. & Hogan, B. *Nature* **317**, 745–748 (1985).
21. Ruddle, F. H. *et al. Cold Spring Harb. Symp. quant. Biol.* **50**, 277–284 (1985).
22. Akam, M. E. *EMBO J.* **2**, 2075–2084 (1983).
23. Levine, M., Hafen, E., Garber, R. L. & Gehring, W. J. *EMBO J.* **2**, 2037–2046 (1983).
24. McGinnis, W., Levine, M., Hafen, E., Kuroiwa, A. & Gehring, W. J. *Nature* **308**, 428–433 (1984).

25. Kornberg, T., Siden, J., O'Farrell, P. & Simon, M. *Cell* **40**, 45–53 (1985).
26. Akam, M. E. & Martinez-Arias, A. *EMBO J.* **4**, 1689–1700 (1985).
27. Harding, K., Wedeen, C., McGinnis, W. & Levine, M. *Science* **229**, 1236–1242 (1985).
28. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
29. McGinnis, W., Hart, C. P., Gehring, W. J. & Ruddle, F. H. *Cell* **38**, 675–680 (1984).
30. Rabin, M. *et al. Nature* **314**, 175–178 (1985).
31. Joyner, A. L. *et al. Nature* **314**, 173–175 (1985).
32. Parnes, J. R. & Seidman, J. G. *Cell* **29**, 661–669 (1982).
33. Kelly, K., Cochran, B. H., Stiles, C. D. & Leder, P. *Cell* **35**, 603–610 (1983).
34. Tilghman, S. M. & Belayew, A. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5254–5257 (1982).
35. Hafen, E. thesis, Univ. Basel (1983).
36. Karch, F. *et al. Cell* **43**, 81–96 (1985).
37. Diwan, S. & Stevens, L. C. *Mouse News Lett.* **51**, 25 (1974).
38. Green, M. C. *Genetic Variants and Strains of the Laboratory Mouse* (Fischer, Stuttgart, 1981).
39. Pravtcheva, D. D., Ruddle, F. H., Ellis, R. W. & Scolnick, E. M. *Somat. Cell Genet.* **9**, 681–686 (1983).
40. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Ratter, W. J. *Biochemistry* **18**, 5294–5299 (1979).
41. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
42. Denhardt, D. T. *Biochem. biophys. Res. Commun.* **29**, 641 (1966).
43. Jolly, D. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 477–481 (1983).
44. Lawrence, J. B. & Singer, R. H. *Nucleic Acids Res.* **13**, 1777–1799 (1985).
45. Gall, J. G. & Pardue, M. L. *Meth. Enzym.* **21**, 470–480 (1971).
46. Hafen, E., Levine, M., Garber, R. L. & Gehring, W. J. *EMBO J.* **2**, 617–623 (1983).
47. Brigati, D. J. *et al. Virology* **126**, 32–50 (1983).
48. Cox, K. H., DeLeon, D. V., Angerer, L. M. & Angerer, R. C. *Dev. Biol.* **101**, 485–502 (1984).

First light from a young star?

Bo Reipurth* & John Bally†

* Copenhagen University Observatory, Oster Voldgade 3,
DK-1350 Copenhagen K, Denmark

† AT&T Bell Laboratories, HOH L-245, Holmdel,
New Jersey 07733, USA

The early phases of stellar evolution are frequently characterized by vigorous interaction between the nascent star and its environment, resulting in the formation of molecular outflows, highly collimated jets, Herbig-Haro objects and small emission or reflection nebulae. A deep survey of such small nebulae around young stars in the Orion molecular clouds has recently been completed¹. We report here the sudden appearance of a highly variable, conical nebula (Object 50) 1.5 arc min south of a $250L_{\odot}$ infrared source in the southern part of the L1641 cloud in Orion. The nebula coincides with the edge of the approaching lobe of a 10^5 -yr-old bipolar molecular outflow. Variability of the infrared source generates light pulses observed to propagate through the surrounding ambient cloud, producing a fluctuating illumination pattern. This emission may represent the first optical radiation to have emerged from this newborn star.

Object 50 is a large (1 arc min wide), luminous and structured nebula located in one of the most opaque regions of the Orion clouds. It is readily visible on a SERC-J Schmidt plate obtained in 1979, but examination of the blue and red Palomar Schmidt plates from 1955 shows nothing at this location. Obviously a highly time-dependent phenomenon is taking place here. A search at $2.2\ \mu\text{m}$ with the ESO 1-m telescope at La Silla, Chile, has shown that Object 50 itself is not directly associated with a

star, but is a reflection nebula of an intense near-infrared source located ~ 1.5 arc min to the north, and coinciding with a faint star-like nebulosity on the Schmidt plate. This source has also been detected by the Infrared Astronomy Satellite (IRAS). Positions measured on the 1979 Schmidt plate are listed in Table 1.

Broad-band CCD (charge-coupled device) images have been obtained of the region with the Danish 1.5-m telescope at La Silla on three occasions in the past few years. Figure 1a shows the region as it appeared on 20 December 1982, through a filter transmitting from $8,500\ \text{\AA}$ to the CCD cut-off at almost $1,100\ \text{\AA}$. Object 50 itself shows a well-defined edge towards the infrared source, an intricate pattern of lighter and darker areas, and a faint tail stretching away from the source. About 1.5 arc min north of Object 50, at the location of the infrared source, two unequal nebulosities (the eastern and western lobes) are visible, separated by an opaque bridge. The position of the infrared source, as determined by IRAS and from ground-based near-infrared observations, coincides with the western lobe to within a few arc s. The eastern lobe is the more luminous and it shows a hint of a very faint structure pointing towards the south-east. A similar image, taken at the same time, but through a broad-band filter centred on $6,500\ \text{\AA}$, shows the same morphology, but the lobes around the infrared source are significantly fainter, demonstrating that there is little $H\alpha$ emission. Figure 1b shows an image in all respects similar to Fig. 1a (same instrumentation, exposure time, filter, seeing, reduction procedure), but obtained two years later, on 9 January 1985. Dramatic changes have occurred in the intervening period. While the western lobe has faded considerably, the faint structures associated with the eastern lobe have evolved into a highly collimated beam, 25 arc s long, with two knots. The length of the beam seen in January 1985 is almost 3 light-months if the scattering surface is located in the plane of the sky. The beam is directed 180° away from the axis connecting the infrared source with the centroid of the redshifted portion of the molecular outflow (see below). The

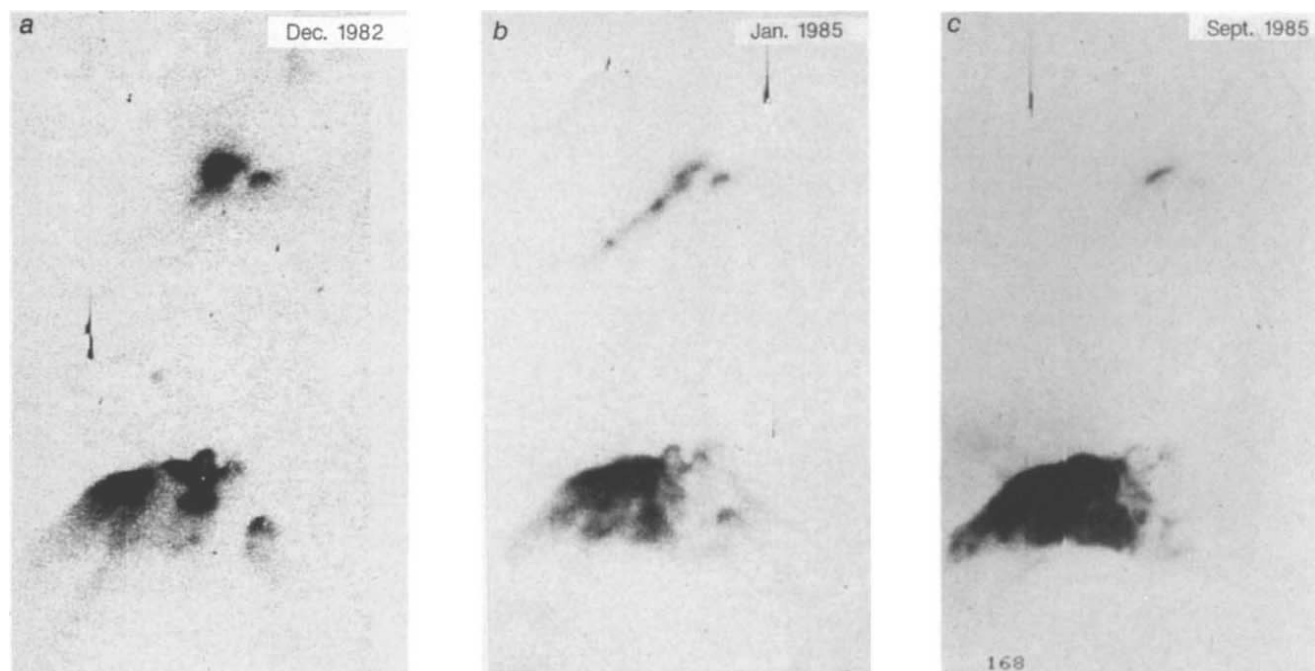


Fig. 1 Three CCD images showing Object 50, taken through broad-band filter (from $\sim 8,500\ \text{\AA}$ to almost $1,100\ \text{\AA}$) with the Danish 1.5-m telescope at La Silla, Chile, on 20 December 1982 (a), 9 January 1985 (b) and 23 September 1985 (c). This nebula appeared after 1955, as nothing is visible on the Palomar Sky Survey plates taken of this region. The field shown is 1.75 arc min in right ascension and 3 arc min in declination; north is at the top and east at the right of each image. The intense and variable infrared source IRAS05380-0728 is situated at the small nebulosity in the upper part of the field. A pulse of light, presumably emitted in late 1984, illuminates the 'jet' seen in the January 1985 image and is observed to propagate 1.5 arc min to the south by September 1985. The faint pattern in the background is residual interference of sky emission lines in the CCD chip.

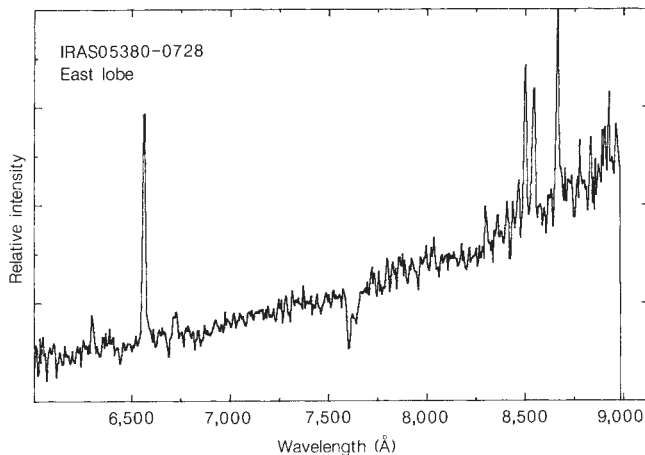


Fig. 2 A low-resolution red spectrum of the nebulosity around the infrared source, obtained on 27 January 1985 with the 4-m telescope at Cerro Tolo by J. A. Graham, using a Boller and Chivens spectrograph and CCD detector. The low extinction inferred from the red continuum indicates that there is much less obstructing material toward the reflection nebula than along our line-of-sight. The spectrum indicates that the hidden star is of low or intermediate mass.

overall morphology of Object 50 is retained, but its luminous pattern has changed completely. Figure 1c shows the region as it appeared on 23 September 1985, displayed with intensity levels roughly similar to Fig. 1a, b; the beam of light extending from the infrared source has all but disappeared, leaving only a faint and diffuse illuminated patch. A faint extended nebulosity is now seen to surround Object 50 and may represent the same pulse of light which produced the beam 9 months earlier. A comparison of the three CCD images shows a gradual fading of the western nebulosity, where the infrared source is located.

Near-infrared photometry of the infrared source IRAS05380-0728, obtained through a 15-arc-s diaphragm on 31 January 1985, is given in Table 2 together with the IRAS fluxes. Comparison with observations made in April and December 1983 reveals considerable variability in the source, with a decrease of $\sim 20\%$ in near-infrared luminosity from 1983 to 1985. For an assumed distance to the L1641 cloud of 460 pc, the luminosity in 1985 between 1.2 and $4.8 \mu\text{m}$ is $\sim 27L_{\odot}$, the IRAS luminosity between 12 and $100 \mu\text{m}$ (without colour correction factors) is $\sim 120L_{\odot}$, and an interpolation between 4.8 and $12 \mu\text{m}$ yields roughly another $100L_{\odot}$. Altogether, the luminosity of the source is $\sim 250L_{\odot}$; this is considerably less than that observed for massive embedded stars, but fairly luminous compared with the average low-mass pre-main-sequence star. Except for the IRAS catalogue, the source is not listed in catalogues of infrared sources. The source FIRSSE101 is only 2.5 arc min away², one-third of the way towards the equally bright source IRAS05375-0731. Since FIRSSE101 was detected with a 10-arc-min beam, it is probably an unresolved blend of the two sources.

We have explored the nature of the infrared source through optical and infrared spectroscopy of the reflected light. Figure 2 shows a low-resolution red spectrum of the luminous eastern

Table 1 Positions measured on the 1979 Schmidt phase

Object	RA (1950)	Dec. (1950)
Object 50	5 h 39 min 05.0 s	$-7^{\circ} 30' 26''$ a
Object 50 IRS (optical position)	5 38 03.4	$-7^{\circ} 29' 04''$ b
Object 50 IRS (IRAS position)	5 38 02.7	$-7^{\circ} 28' 59''$

IRS, infrared source; a, approximate central position; b, brightest eastern lobe.

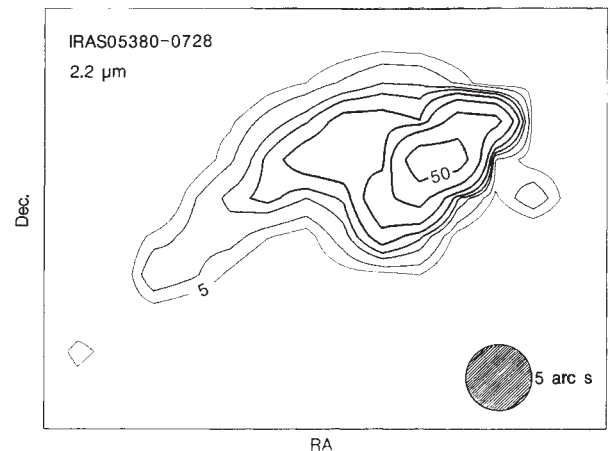


Fig. 3 A 2.2- μm map of the infrared source and its immediate surroundings, obtained with a 5-arc-s diaphragm through the ESO 1-m telescope at La Silla on 9 February 1985. The contour levels are at 5, 7, 10, 12, 15, 20, 30 and 50 (arbitrary units). The source is clearly extended along the axis determined by the source and the centre of the redshifted CO lobe. The contours of the infrared reflection nebula correspond very well to the optical appearance of the object at the same time. The ruled circle shows the size of the aperture stop used to make the map; it corresponds to a 5-arc-s diameter circle on the sky.

lobe east of the infrared source, obtained on 27 January 1985 at the CTIO 4-m telescope with a Boller and Chivens spectrograph and a CCD. The spectrum shows a slightly reddened continuum with purely terrestrial absorption features and only a few emission lines: mainly $H\alpha$ and the calcium triplet, as well as weak line of [O I] and [S II]. This appearance is very different from that of the more massive pre-main-sequence stars⁴. In accordance with the luminosity estimate, the spectroscopic

Table 2 Infrared observations

	J	H	K	L	M
31 January 1985 (mag.)	13.8 (0.25)	10.74 (0.05)	8.33 (0.01)	3.72 (0.01)	2.47 (0.03)
	12 μm	25 μm	60 μm	100 μm	
IRAS (Jy)	28.3	89.6	186.6	225.0	

evidence suggests that the infrared source and the source of illumination of Object 50 is a low- or intermediate-mass pre-main-sequence star. The sudden appearance of Object 50 implies an abrupt brightening of the infrared source reminiscent of FU Ori eruptions⁵. However, following an outburst, FU Ori stars do not show $H\alpha$ emission, at least not strong enough to be visible in low-resolution spectra. Moreover, the 2.0–2.2- μm wavelength range in these objects is characterized by a suppression caused by the 1.9- μm H_2O absorption band. Our 2.0–2.5- μm circular variable filter (CVF) spectrum shows no such absorption, only a red continuum with no lines. A 2.8–4.2- μm CVF spectrum shows a pronounced 3.07- μm ice absorption band.

We obtained a map of the source at 2.2 μm on 9 February 1985, with the ESO 1-m telescope and a 5-arc-s diaphragm, by taking data at 3-arc-s intervals in right ascension and declination (Fig. 3). The infrared source is extended in a NW-SE direction, with a structure closely mimicking the optical beam seen in Fig. 1b. The linear optical and near-infrared feature is seen to emanate from the eastern side of the stellar source which lies at or slightly west of the western optical knot. The axis defined by the linear feature points directly toward the centre of the redshifted lobe of the bipolar CO outflow located to the north-west of the infrared source. The prototype of such infrared nebulae is located in Chamaeleon^{6,7}.

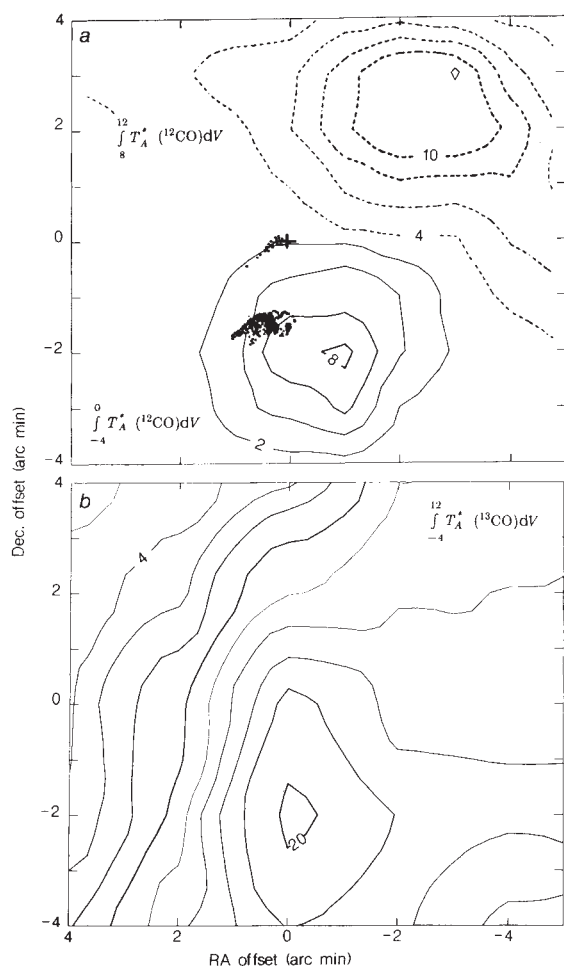


Fig. 4 *a*, A ^{12}CO map made with the 7-m millimetre-wave antenna at Crawford Hill in New Jersey, showing the redshifted (broken contours) and blueshifted (solid contours) lobes of the bipolar flow associated with Object 50. Contour interval: 2 K km s^{-1} . The + symbol at (0, 0) marks the position of the infrared source. The position of the optical nebula is indicated by the dotted pattern. *b*, A ^{13}CO map showing the distribution of total emissivity (which is roughly proportional to the column density of molecular gas) in the region surrounding Object 50. Contour interval: 2 K km s^{-1} . T_A^* is the observed antenna temperature corrected for atmospheric attenuation. V is the velocity (in km s^{-1}) with respect to the local standard of rest.

A 10×10 -arc-min field was mapped in the ^{12}CO and ^{13}CO species with the AT&T Bell Laboratories 7-m millimetre-wave antenna located in Holmdel, New Jersey. The beam size of this antenna at 110 GHz is 100 arc s. An 8-arc-min-long, bent bipolar outflow (Fig. 4*a*) is centred near the infrared source. The velocity of the outflow is low; the full width of the high-velocity emission at the 0.2-K level is only $\sim 16 \text{ km s}^{-1}$. As we do not know the inclination of the source with respect to the line-of-sight, the true flow velocity is impossible to determine; however, a reasonable guess is that it is in the range $4\text{--}8 \text{ km s}^{-1}$. Using the observed isotopic intensity ratio, $^{12}\text{CO}/^{13}\text{CO} = 20$ in the wings, we estimate the observed flow mass to be $\sim 2.5 M_\odot$. About this much mass may be present at velocities hidden by the cloud core and is thus unobservable. Therefore, a good estimate for the total flow mass is $5 M_\odot$. The flow dynamic lifetime is $R/V_{\text{flow}} \approx 10^5 \text{ yr}$, making this one of the older bipolar outflows known⁸. The mechanical luminosity in the flow is $0.6 L_\odot$, the energy of the flow is $2 \times 10^{45} \text{ erg}$, and the momentum is $30 M_\odot \text{ km s}^{-1}$. These values are characteristic of a highly evolved outflow associated with a star of intermediate luminosity.

The ^{13}CO observations (Fig. 4*b*) show that the infrared source is located at the northern end of a high-column-density region centred at V_{LSR} (velocity with respect to the local standard of

rest) $= 4 \text{ km s}^{-1}$. We believe that the blueshifted portion of the CO outflow has been deflected to the west by the north-south ridge of gas. The optical nebulosity is situated at the extreme eastern edge of the blueshifted portion of the bipolar flow and is possibly a protrusion of high-density gas directly illuminated by the star through the cavity excavated by the outflow. This feature may be part of the structure responsible for the deflection of the blueshifted flow. Evidently this is the first place where the flow has punched a hole in the ambient foreground cloud, allowing scattered light from the infrared source to escape towards our line-of-sight.

The abrupt appearance of Object 50 within the past 30 years may indicate a sudden brightening of the infrared source, similar to an FU-Orionis eruption. Variability of reflection nebulae has been observed before, notably in Hubble's variable nebula (NGC 2261)⁹, PV Cep^{10,11}, and R CrA (ref. 12 and J. A. Graham and A. C. Phillips, in preparation), but not on the spectacular scale reported here. We believe the underlying mechanism operating both in the present region and in the above objects is time variability of the light output from a young embedded star, possibly modulated by the movement of opaque masses very close to the star, causing patterns of light and shadow to move over the surrounding cloud. The extended cavities through which the light is transmitted to the reflecting surfaces have been produced recently by a bipolar outflow from the newborn star. This is probably a common phenomenon for all stars that are emerging from their parental clouds. In view of the sudden appearance of Object 50, we may here be seeing the first light emerging from this young star.

We thank John Graham for the optical spectrum, Patrice Bouchet for giving us observing time to obtain the $2.2\text{-}\mu\text{m}$ map, and Leif Hansen for a CCD image. This study was supported in part by the Danish Astronomical Board and the Danish Space Agency.

Received 9 December 1985; accepted 6 February 1986.

1. Reipurth, B. *Astr. Astrophys. Suppl.* **61**, 319–330 (1985).
2. Price, S. D., Murdoch, L. P. & Shivanadan, K. *Far Infrared Sky Survey Experiment. Final Report* AFGL-TR-83-0055 (NASA, Washington, DC, 1983).
3. McGregor, P. J., Persson, S. E. & Cohen, J. G. *Astrophys. J.* **286**, 609–629 (1984).
4. Herbig, G. H. & Soderblom, D. R. *Astrophys. J.* **242**, 628–637 (1980).
5. Herbig, G. H. *Astrophys. J.* **217**, 693–715 (1977).
6. Schwartz, R. D. & Heinze, K. G. *Astr. J.* **88**, 1665–1669 (1983).
7. Cohen, N. & Schwartz, R. D. *Astr. J.* **89**, 277–279 (1984).
8. Bally, J. & Lada, C. J. *Astrophys. J.* **265**, 824–847 (1983).
9. Hubble, E. P. *Astrophys. J.* **44**, 190–197 (1916).
10. Cohen, M., Kuhl, L. V. & Harlan, E. A. *Astrophys. J.* **215**, L127–L129 (1977).
11. Cohen, M., Kuhl, L. V., Harlan, A. E. & Spinrad, H. *Astrophys. J.* **245**, 920–926 (1981).
12. Knox-Shaw, H. *Mon. Not. R. astr. Soc.* **76**, 646–647 (1916).

Occurrence of liquid-crystalline mesophases in microemulsion dispersions

J. Tabony

CEA-IRDI-DESICP, Département de Physico-Chimie, Centre d'Etudes Nucléaires de Saclay, 91191 Gif-sur-Yvette, Cedex, France

Although microemulsion dispersions have compositions which are related to lyotropic liquid crystals, their possible structures are not normally considered to be crystalline. As described here, small-angle neutron diffraction spectra have been obtained from some oil-water microemulsion dispersions, and display peaks which correspond to lamellar, hexagonal and cubic arrangements, similar to those observed in lyotropic liquid crystals. The formation of these liquid crystal mesophases (mesomorphic phases) seems to be a general feature of microemulsion compositions and not an isolated example.

Surfactant molecules in aqueous solution self-aggregate in various ways^{1–4}. At surfactant concentrations of less than about 20–30 wt%, micellar solutions are formed, which comprise small

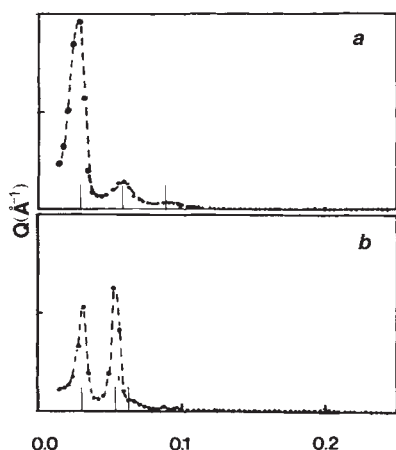


Fig. 1 Neutron small-angle diffraction patterns from microemulsions with the compositions given in the text. *a*, Composition A; *b*, composition B. The vertical lines indicate the expected positions⁹ for *a*, a lamellar structure with 185-Å spacing and *b*, a hexagonal structure with a 200-Å unit cell.

quasi-spherical aggregates, with a radius approximately the size of the surfactant chain length^{3,4}. Little or no long-range order exists between these micelles. At higher surfactant concentrations the micelles often increase in size, becoming rods or plates; long-range order is established and the well known lamellar, hexagonal and cubic liquid crystalline mesophases occur^{1,2}. Solubilization of hydrocarbons in these surfactant-water binary mixtures is often limited, but is considerably enhanced by the addition of a co-surfactant such as a short-chain alcohol like pentanol or butanol.

The resulting oil and water dispersions, commonly termed microemulsions⁵, are frequently fluid, transparent and optically isotropic. Often such microemulsion dispersions can be formulated over a wide range of oil and water volume fractions. For small concentrations of oil or water the structure has frequently been found to be a dispersion of oil droplets in water or water droplets in oil⁶. The droplets are stabilized by an interfacial film of surfactant and co-surfactant and may be considered to be swollen micelles.

Microemulsion dispersions are not normally thought to form liquid-crystalline mesophases. However, since in surfactant-water binary mixtures micelles precede the appearance of liquid-crystalline phases, and since some microemulsion dispersions resemble swollen micelles, then under some circumstances, such as an increased volume fraction of swollen micelles, the oil or water droplets may rearrange in such a way that liquid-crystalline materials result. Recently, a neutron diffraction pattern consistent with a cubic structure has been obtained from a microemulsion containing approximately equal volumes of oil and water⁷. Here I present neutron diffraction spectra from microemulsion dispersions which are characteristic of lamellar and hexagonal structures. Unlike the cubic phase mentioned above, these liquid-crystalline phases occur for low as well as high volume fractions of dispersed oil or water. These liquid crystalline materials differ from ordinary surfactant-water lyotropic liquid crystals in that the surfactant concentrations are much lower (10 wt% or less) than those normally used (~50 wt%). This results in lattice spacings (200–1,000 Å) considerably larger than those found in surfactant-water binary systems (20–100 Å). One way of accounting for this difference in behaviour is the following: if in binary surfactant-water liquid crystals, the liquid-crystalline phases are derived from small surfactant micelles (radius ~20 Å), then in microemulsions they will arise from the larger swollen micelles (radius 100–200 Å). Hence in the latter case not only will the crystal spacings be much greater, but because the bulk of the swollen micelle is comprised of either oil or water, the amount of surfactant in the dispersion

will be much less than for the binary surfactant-water case. Moreover, as the bulk of the surfactant is situated at the oil-water interface, the ratio of the quantities of surfactant to oil or surfactant to water will control the dimensions of the crystal spacings. Such a simplistic explanation does not, however, explain why such liquid crystals arise. They may no longer be microemulsions in the strict sense of the word, and the use of this term may be misleading. Here I shall continue to call them liquid crystal microemulsions, although oil-water liquid crystals might be a more appropriate name.

Figure 1 shows neutron diffraction spectra obtained from two liquid crystal microemulsions, which display lamellar and hexagonal structures. The spectra were obtained on the small-angle neutron diffractometer, D16, at the Institut Laue-Langevin, Grenoble⁸.

The microemulsion compositions were: A (lamellar), tetradecyl trimethylammonium bromide (0.58 g), pentanol (0.65 ml), cyclohexane (0.35 ml), D₂O (8 ml); B (hexagonal), tetradecyl trimethylammonium bromide (1.42 g), pentanol (0.60 ml), cyclohexane (4.0 ml), D₂O (4.0 ml). So as to allow comparison, the composition of the cubic phase already reported⁷ was tetradecyl trimethylammonium bromide (2.1 g), butanol (1.0 ml), octane (4.0 ml), D₂O (4.0 ml).

The spectrum from composition A (Fig. 1*a*) shows three equidistant Bragg peaks corresponding to a lamellar⁹ structure with a lattice spacing of 185 Å. Although the dispersion contains more than 85% water it is both viscous and optically birefringent. The spectrum from composition B (Fig. 1*b*) shows Bragg peaks for values of the momentum transfer (*Q*) in approximately the ratio 1:√3:2. This is consistent with a hexagonal arrangement with a unit cell of 200 Å. Like the lamellar phase, the dispersion is both viscous and birefringent. As in ordinary lyotropic liquid crystals, the viscosity increases in the order lamellar, hexagonal, cubic: the lamellar phase flows slowly, whereas the cubic phase appears solid, flowing only over a period of days or weeks. None of these materials appears to behave as a newtonian fluid; they form spontaneously on mixing and show no signs of separation either after storage for more than a year or on centrifugation. Hence, they appear to be thermodynamically stable, and probably do not consist of a dispersion of small liquid-crystal particles. If the latter were the case, the resulting Bragg spacings would be much smaller than those observed.

The structure of these phases has not yet been fully elucidated. The lamellar phase probably comprises thin layers of cyclohexane sandwiched between an interfacial film of surfactant and co-surfactant. The amount of cyclohexane in composition A corresponds to approximately one monolayer; given the length of the surfactant molecule, this suggests that the oil-soap sandwich should be approximately 30–40 Å thick. This should be compared with the overall spacing of 185 Å. These mesophases occur over a relatively restricted range of alcohol concentrations. For the lamellar phase this corresponds to a molar ratio of alcohol to surfactant of ~3:1. This observation suggests that in this case the interfacial layer is comprised of a two-dimensional hexagonal arrangement of alcohol molecules with a surfactant molecule in the centre. The proton NMR spectra consist of relatively sharp lines similar to those observed in ordinary lyotropic liquid crystals; this shows that the different molecular components are in a mobile liquid state. It thus seems likely that the interfacial layer of surfactant and co-surfactant has considerable dynamic disorder. Although the lamellar composition (A) contains only a small amount of cyclohexane, much higher cyclohexane concentrations may be used. Lamellar phases have been observed in compositions having equal volumes of oil and water, and even oil-based compositions. It thus appears that the thickness of the oil and/or water layers can be varied almost at will.

In the phase diagram, the lamellar phase of composition A adjoins that of a dispersion of swollen micelles, and changes in the alcohol or surfactant concentration result in a transition to

such a structure. This suggests that the lamellar phase is closely related to a dispersion of droplets. A transition from a dispersion of droplets to a dispersion of plates could result in a lamellar nematic-type phase. The driving force for the transition is clearly related to the alcohol content of the dispersion. Following the interfacial theory (R) presented by Winsor¹⁰, it can be argued that progressive intercalation of alcohol amongst the paraffinic surfactant chains will tend to reduce the interfacial curvature in the swollen micelle until flat regions are formed which subsequently order in a lamellar manner. This is supported by the observation that in the hexagonal and cubic phases, which have locally non-zero interfacial curvatures, the alcohol-to-surfactant ratio is less than in the lamellar phase.

Although all of the examples presented contain the same surfactant, I have also formulated liquid crystal microemulsions with other surfactants and co-surfactants. Their occurrence hence seems to be a general feature of microemulsion phase behaviour and not an isolated example.

I thank the Institut Laue-Langevin, Grenoble, for providing the neutron beam time, and S. Wilson and L. Braganza for their assistance with the D16 diffractometer.

Received 9 September 1985; accepted 21 January 1986.

1. Elkwall, P. in *Advances in Liquid Crystals* Vol. 1 (ed. Brown, G. H.) Ch. 1 (Academic, New York, 1971).
2. Tiddy, G. J. T. *Phys. Rep.* **57**, 1–46 (1980).
3. Wennerstrom, H. & Lindman, B. *Phys. Rep.* **52**, 1–86 (1979).
4. Lindman, B. & Wennerstrom, H. *Topics Curr. Chem.* **87**, 1–84 (1980).
5. Prince, L. M. *Microemulsion, Theory and Practice* (Academic, New York, 1977).
6. Cebula, D. J., Harding, L., Ottewill, R. H. & Pusey, P. N. *Colloid Polym. Sci.* **258**, 973–976 (1980).
7. Tabony, J. *Nature* **319**, 400 (1986).
8. *Neutron Beam Facilities Available for Users* (Institut Laue-Langevin, Grenoble, 1981).
9. Fontell, K. in *Liquid Crystals and Plastic Crystals* Vol. 2 (eds Gray, G. W. & Winsor, P. A.) Ch. 4 (Wiley, New York, 1974).
10. Winsor, P. A. in *Liquid Crystals and Plastic Crystals* Vol. 1 (eds Gray, G. W. & Winsor, P. A.) Ch. 5 (Wiley, New York, 1974).

Phase behaviour of concentrated suspensions of nearly hard colloidal spheres

P. N. Pusey & W. van Megen*

Royal Signals and Radar Establishment, St Andrews Road, Malvern WR14 3PS, UK

Suspensions of spherical colloidal particles in a liquid show a fascinating variety of phase behaviour which can mimic that of simple atomic liquids and solids. 'Colloidal fluids'^{1–4}, in which there are significant short-range correlations between the positions of neighbouring particles, and 'colloidal crystals'^{5–7}, which have long-range spatial order, have been investigated extensively. We report here a detailed study of the phase diagram of suspensions of colloidal spheres which interact through a steep repulsive potential. With increasing particle concentration we observed a progression from colloidal fluid, to fluid and crystal phases in coexistence, to fully crystallized samples. At the highest concentrations we obtained very viscous samples in which full crystallization had not occurred after several months and in which the particles appeared to be arranged as an amorphous 'colloidal glass'. The empirical phase diagram can be reproduced reasonably well by an effective hard-sphere model. The observation of the colloidal glass phase is interesting both in itself and because of possible relevance to the manufacture of high-strength ceramics⁸.

The particles studied were colloidal polymethylmethacrylate (PMMA), stabilized sterically by poly-12-hydroxystearic acid⁹. Electron microscopy, light scattering and crystallography showed the radius of the PMMA cores to be 305 ± 10 nm and

the polydispersity (standard deviation of the particle size distribution divided by the mean size) to be about 0.05. The thickness of the stabilizer layer was 10–20 nm, a small fraction of the particle radius. Thus the repulsive potential arising from interpenetration of the polymer coatings of different particles was relatively steep. The suspension medium was a mixture of decalin and carbon disulphide in volume ratio 2.66:1, chosen to match closely the refractive index of the particles, ~ 1.51 (ref. 10). This provided nearly transparent samples suitable for light scattering measurements as well as for visual observation of processes occurring in the bulk of the samples. Furthermore, index matching is expected to minimize interparticle attractions due to van der Waals forces. Ten samples were prepared from a stock solution of known PMMA weight fraction (determined by drying). They were concentrated by low-speed centrifugation to form a dense sediment, followed by removal of a weighed amount of clear supernatant liquid. Slow tumbling of the samples then redispersed the particles effectively. The fractional volumes ϕ_c of the samples occupied by PMMA cores were calculated using literature values for the densities of PMMA and the liquids.

After extensive tumbling which, we assume, left the particles positionally randomized (that is, showing only local short-range order), samples 2–10 were set up as shown in Fig. 1a and were illuminated obliquely from behind by a broad beam of white light. The most dilute sample (2, on the extreme right of Fig. 1a), with $\phi_c = 0.393$, showed no macroscopically observable change with time. Earlier work^{10–12} has established that at (and below) this concentration the particles are spatially arranged much like atoms in a dense liquid, exhibiting considerable short-range positional ordering. Although hindered by their neighbours, the particles remain able to diffuse through this colloidal fluid under the influence of brownian motion. After times ranging from minutes to hours, small, Bragg-reflecting crystallites formed homogeneously throughout the volumes of samples 3–7. In samples 3–5 the crystallites settled under gravity within a day to form well-defined boundaries between coexisting polycrystalline and fluid phases. Samples 6 and 7 remained filled with small compact crystallites. In these two samples the state of lowest free energy is evidently a colloidal crystal which grows, by diffusion of the particles, from the randomized metastable state formed by the tumbling process. Once crystallization is complete, the particle motions are largely limited to local brownian excursions centred on sites in the regular crystalline array. The crystal structure was determined by light-scattering crystallography to be face-centred cubic¹³. The coexistence of colloidal fluid and crystal observed in samples 3–5 is analogous to the coexistence of a simple liquid and solid (such as ice and water) at a first-order phase transition. Sample 8 crystallized heterogeneously, starting from the meniscus and cell walls, to form relatively large irregular crystals. The most concentrated ($\phi_c \geq 0.50$), highly viscous samples 9 and 10 exhibited only partial heterogeneous crystallization even when left undisturbed for several months. Although the crystalline state is, presumably, still thermodynamically preferred for these samples, it appears that the concentration is so high that particle diffusion is hindered to the point where crystals do not form on this timescale and the suspensions remain in the metastable amorphous phase created by the tumbling. The different results of homogeneous and heterogeneous nucleation are shown clearly in Fig. 1b.

The observed phase behaviour is summarized in Fig. 2. In the coexistence region we plot the fraction of the sample volume occupied by the crystalline phase against ϕ_c . Not surprisingly, we find a linear (lever-rule) dependence which, when extrapolated to 0 and 100%, provides 'freezing' and 'melting' concentrations.

While there is evidence¹³ that, due to interpenetration of their stabilizing polymer coatings, the interaction between the particles is slightly 'soft', it seems simplest to discuss our findings in terms of an effective hard-sphere model; such models for simple

* Permanent address: Department of Applied Physics, Royal Melbourne Institute of Technology, Melbourne, Australia.

Fig. 1 *a*, Nearly-hard-sphere suspensions, 4 days after tumbling, illuminated obliquely from behind by white light. The samples are contained in cells of cross-section $1\text{ cm} \times 1\text{ cm}$ and are numbered 2–10, in our nomenclature, from the right in order of increasing concentration. Actual concentrations are shown in Fig. 2. Immediately after tumbling, all the samples appeared featureless (see top of sample 2 or bottom of sample 9), reflecting randomization of the particle positions. The text describes the dynamic processes which occurred over 4 days and led to the structures depicted in the photograph. The scattering geometry is such that the first peak in the structure factor of the amorphous phases and the first Bragg reflection of the crystals occur in the red at the right of the picture and in the blue-green at the left. Some overall gravitational settling of the particles has occurred, leading to a small layer of crystals at the bottom of sample 2 (which, it should be emphasized, showed no crystallization in the bulk) and compaction of the crystallites at the bottom of samples 3–7. *b*, Close-up of samples 7–9 under the same conditions as in *a*. Crystallization nucleated homogeneously in sample 7 but heterogeneously in sample 8. The lower part of sample 9 shows clearly the amorphous glassy phase which did not crystallize over several months.



liquids have been extensively studied, both theoretically and by computer simulation^{14–17}. Thus we assume that the core volume fraction $\phi_C^E = 0.407$, at which freezing starts (see Fig. 2), is equivalent to an effective volume fraction 0.494, the freezing concentration for hard spheres¹⁶. This implies a hard-sphere interaction radius about 20 nm (roughly the thickness of the stabilizer layer) greater than the core radius, $\sim 305\text{ nm}$. We therefore include in Fig. 2 a scale appropriate to the effective hard-sphere volume fraction, defined by $\phi_E = (0.494/0.407) \phi_C$. With this scaling the melting density $\phi_E^M = 0.536$ for the colloidal suspension is slightly smaller than the value 0.545 expected for the equivalent hard spheres¹⁶. Such a slight difference between the observed coexistence region and that expected for hard spheres could be caused by several factors; for example, softness

of the interparticle potential¹⁷, slight interparticle attractions¹⁷, particle polydispersity¹⁸, or, very likely, a combination of these. The effective volume fraction of the most concentrated glassy sample, 10, is close to the value $\phi_E = 0.637$ expected for the Bernal (random close-packed) hard-sphere glass^{14,19,20}.

We conclude that the colloidal suspensions described above should prove useful model 'nearly-hard-sphere' systems for further experimentation. In particular, the rate of transport in these suspensions is many orders of magnitude smaller than that in simple molecular systems, a property which should simplify the study of some kinetic processes²¹. We have already made a number of other measurements, discussed elsewhere¹³. These include: the kinetics of crystallization, light crystallography, rheological measurements in the metastable fluid and

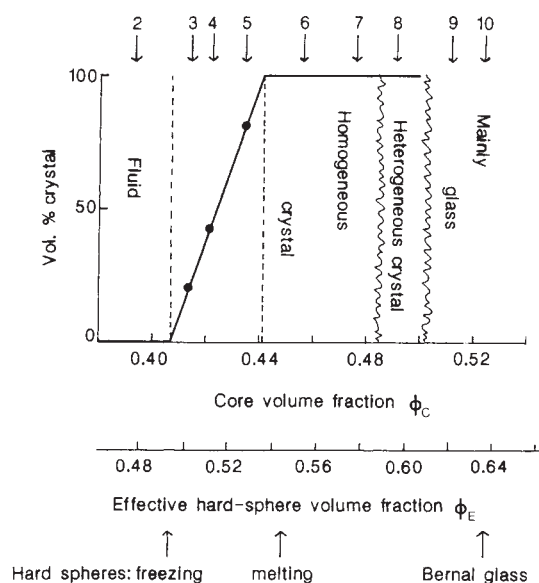


Fig. 2 'Phase-diagram' of the samples shown in Fig. 1. Arrows at top indicate the sample concentrations. The two horizontal axes indicate the measured volume fraction of PMMA cores and the effective hard-sphere volume fraction defined in the text. Arrows at the bottom indicate the volume fractions, obtained from computer simulations, for freezing, melting and random close packing ('Bernal glass') of hard spheres.

glassy phases, and measurements by conventional and dynamic light scattering of structure factors and particle diffusion in the fluid, crystal and glassy phases. The discovery of a glass composed of (nearly) equal-sized spheres is especially interesting since, although there has recently been considerable theoretical work (for example, refs 22–25) and computer simulations^{14,15} on such model glasses, real glasses composed of spherical units are rare. (Colloidal glasses observed up to now have contained mixtures of spheres of different sizes²⁶.) Finally we note current interest⁸ in manufacturing industrial ceramics from 'green bodies' composed of particles of uniform size packed densely either at random (a colloidal glass) or in a colloidal crystal; the present work shows how both these structures may be achieved.

We thank Professor R. H. Ottewill and his group at Bristol University for providing the PMMA particles complete with electron-microscopic characterization; Dr L. V. Woodcock for a valuable discussion; and Mr L. Clarke for the photography.

Received 13 November 1985; accepted 5 February 1986.

1. Brown, J. C., Pusey, P. N., Goodwin, J. W. & Ottewill, R. H. *J. Phys. A* **8**, 664–682 (1975).
2. Dickinson, E. A. *Rep. R. Soc. Chem. C*, 3–37 (1983).
3. van Megen, W. & Snook, I. *Adv. Colloid Interface Sci.* **21**, 119–194 (1984).
4. Edwards, J., Everett, D. H., O'Sullivan, T., Pangalou, I. & Vincent, B. *JCS Faraday Trans. I*, **80**, 2599–2607 (1984).
5. Kose, A. & Hachisu, S. *J. Colloid Interface Sci.* **46**, 460–469 (1974).
6. Clark, N. A., Hurd, A. J. & Ackerson, B. J. *Nature* **281**, 57–60 (1979).
7. Pieranski, P. *Contemp. Phys.* **24**, 25–73 (1983).
8. Calvert, P. *Nature* **317**, 201 (1985).
9. Antl, L. *et al. Colloids and Surfaces* (in the press).
10. van Megen, W., Ottewill, R. H., Owens, S. M. & Pusey, P. N. *J. chem. Phys.* **82**, 508–515 (1985).
11. Pusey, P. N. & van Megen, W. *J. Phys., Paris* **44**, 285–291 (1983).
12. Vrij, A. *et al. Faraday Discuss. chem. Soc.* **76**, 19–35 (1983).
13. Pusey, P. N. & van Megen, W. in *Proc. Symp. Physics of Complex and Supermolecular Fluids* (ed. Safran, S. A.) (Wiley, New York, in the press).
14. Woodcock, L. V. *Ann. N.Y. Acad. Sci.* **37**, 274–298 (1981).
15. Angell, C. A., Clarke, J. H. R. & Woodcock, L. V. *Adv. chem. Phys.* **48**, 397–453 (1981).
16. Hoover, W. G. & Ree, F. H. *J. chem. Phys.* **49**, 3609–3617 (1968).
17. Hansen, J. P. & McDonald, I. R. *Theory of Simple Liquids* (Academic, New York, 1976).
18. Dickinson, E. & Parker, R. J. *J. Phys., Paris* **46**, L229–L232 (1985).
19. Bernal, J. D. & Mason, J. *Nature* **188**, 910 (1960).
20. Bernal, J. D. *Proc. R. Soc. A* **280**, 299–322 (1964).
21. Hanley, H. J. M., Rainwater, J. C., Clark, N. A. & Ackerson, B. J. *J. chem. Phys.* **79**, 4448–4458 (1983).
22. Leutheusser, E. *Phys. Rev. A* **29**, 2765–2773 (1984).
23. Das, S. P., Mazenko, G. F., Ramaswamy, S. & Toner, J. *J. Phys. Rev. Lett.* **54**, 118–121 (1985).
24. Singh, Y., Stoessel, J. P. & Wolynes, P. G. *Phys. Rev. Lett.* **54**, 1059–1062 (1985).
25. Ullo, J. J. & Yip, S. *Phys. Rev. Lett.* **54**, 1509–1512 (1985).
26. Lindsay, H. M. & Chaikin, P. M. *J. chem. Phys.* **76**, 3774–3781 (1982).

Bubble raft model for indentation with adhesion

J. M. Georges, G. Meille, J. L. Loubet & A. M. Tolen

Laboratoire de Technologie des Surfaces, Ecole Centrale de Lyon, 36 avenue Guy de Collongue, CNRS UA 85, BP 163, 69131 Ecully Cedex, France

Indentation hardness tests are now widely used to measure the mechanical properties of solid surfaces^{1–3}. Recent developments of this technique^{4,5} permit the analysis of the outermost 10 nm of materials. Experimental and theoretical questions arise regarding the physical and mechanical processes involved in such small indentations. We describe here an indentation experiment on a microscopic scale, using soap bubbles blown onto a water surface. Bubble rafts provide a simple two-dimensional model for indentation behaviour; as for other materials, their behaviour is governed by two principal attraction–repulsion forces⁶, and by geometrical constraints. A crystalline two-dimensional lattice is obtained by using bubbles of uniform size^{7–9}, whereas bubbles of two sizes give an amorphous structure^{10,11}. Indentation can be represented by the contact between a triangular crystalline raft and a rectangular crystalline raft bordered by an amorphous layer. The flow of the materials, which is dependent on both adhesion and the force between the two rafts, can be analysed during the experiment.

The experimental configuration is shown in Fig. 1. We used a rectangular glass trough, approximately 60 × 30 cm and 15 cm deep, containing a soap solution of surface tension $\gamma = 6.9 \times 10^{-2} \text{ N m}^{-1}$. Air at constant pressure was blown through short capillaries with 50- μm and 100- μm bores in order to produce uniform bubbles. The bubble diameter depends on the nozzle size and is independent of the distance of the jet below

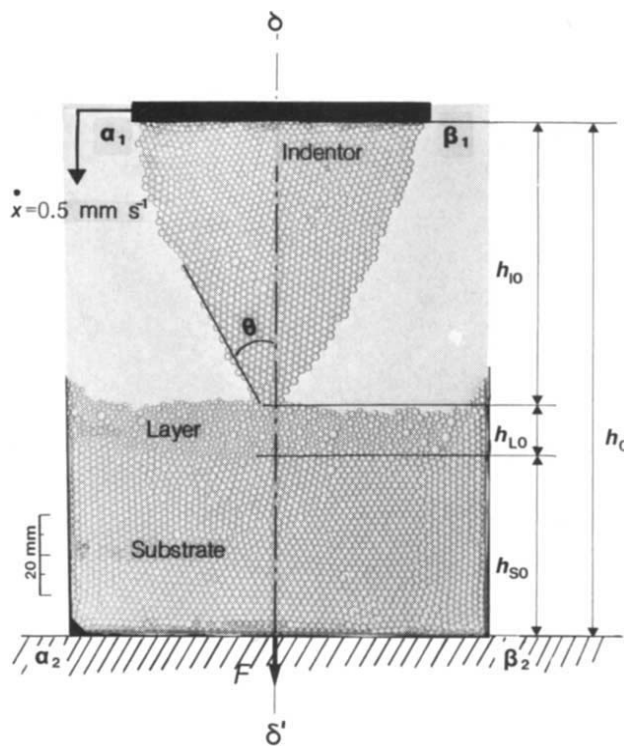


Fig. 1 Bubble rafts used to model the indentation process. Two-dimensional rafts containing bubbles of 2.3 mm and 1.4 mm diameter are blown onto a water surface. The displacement to frame $\alpha_1\beta_1$ parallel to frame $\alpha_2\beta_2$ causes indentation to take place. The force F is measured in the plane perpendicular to frame $\alpha_2\beta_2$.

the surface and the pressure. A raft with uniform bubble diameter (2.3 mm) gives a crystalline structure; a raft with a 40/60 mixture of two diameters (2.3 and 1.4 mm) gives an amorphous structure.

The indenter is a triangular sheet of large bubbles with tip angle $2\theta = 60^\circ$ which sticks to the frame $\alpha_1\beta_1$ (Fig. 1). The indented material is a layer of a mixture of two sizes of bubble (thickness h_{LO}) on a substrate which is a rectangular sheet of large bubbles (thickness h_{SO}) attached to frame $\alpha_2\beta_2$. Frame $\alpha_1\beta_1$ is moved at constant low speed (0.5 mm s^{-1}) parallel to

frame $\alpha_2\beta_2$, causing an indentation in which a displacement, rather than a load, is imposed. As indentation proceeds, the normal force F transmitted to $\alpha_2\beta_2$ is continuously recorded, giving the loading curve (see Fig. 3). The movement is then reversed to obtain the unloading curve.

An enlarged motion picture of the experiment reveals some important aspects of the indentation process (Fig. 2). As the indenter approaches the surface, adhesion occurs between the indenter tip and the amorphous layer. Attractive forces between the two rafts appear well before the two touch, at a distance of about 5 mm. On moving $\alpha_1\beta_1$ further, repulsive forces appear, and the first displacements occur in the amorphous layer. Then the indenter sinks into this layer, which clearly undergoes considerable plastic flow. In spite of local slips, the substrate does not yield perceptibly under the thrust until the penetration reaches a point where only a small amount of amorphous material (a thickness of 5 or 6 bubbles) is trapped in the interface. It is essential to note that, however far we pushed the indenter into the layer and substrate (3, Fig. 2), we could never penetrate the amorphous layer.

On unloading, the amorphous material sticks to the indenter tip so that tensile flow results in contraction and finally rupture in the glassy structures. Therefore, the 'surface' after testing does not have a regular triangular deformation but a hollow imprint with a protruberance in the middle. The indenter tip has an adhered layer, a fact that should be kept in mind when performing repeated microindentation measurements.

Figure 3 shows the loading and unloading curves, obtained by plotting the load F against the displacement $x = h_0 - h$, where h is the current distance between the two frames. Repulsive and compressive forces lie above the abscissa; attractive and tensile forces lie below it. The arrowheads indicate the progression of the experiment. As loading begins (A), the initial force is attractive, due to adhesion between the indenter tip and the amorphous layer. With progressive deformation, adhesion becomes less important, and F increases broadly proportionally to x ; superimposed on this linear trend are some sudden instabilities caused by grain slips and dislocation movements in the crystalline rafts. Towards the end of the loading the force rises steeply, thus confirming the earlier assumption that loading consists of

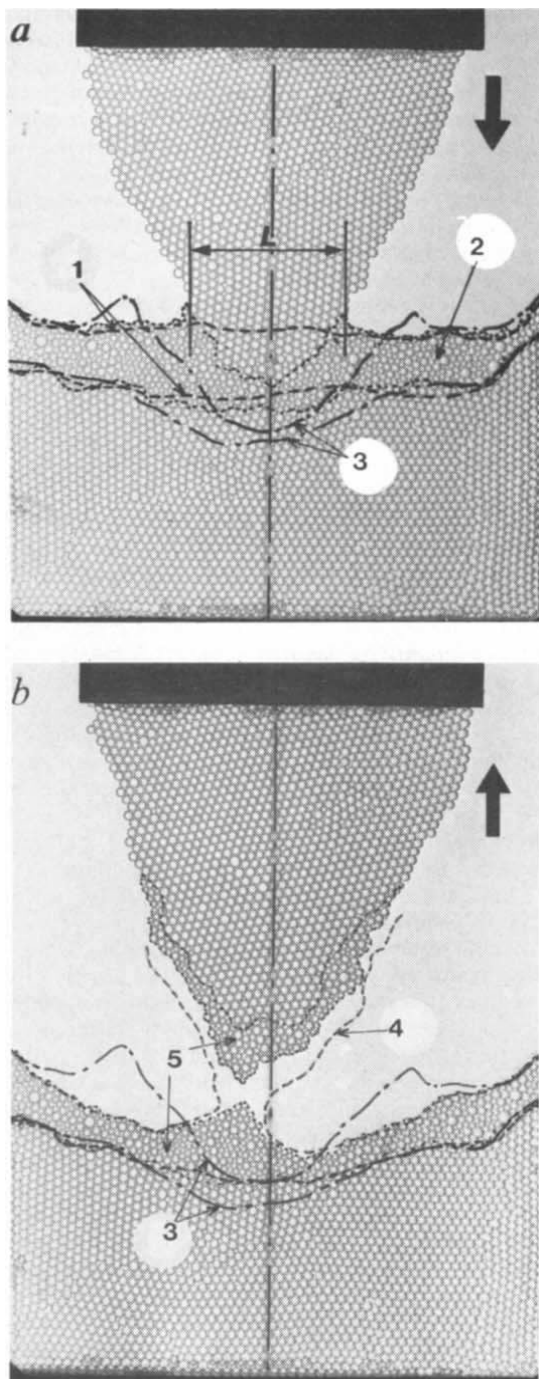


Fig. 2 Examples of deformation during loading (a) and unloading (b). The boundaries of the amorphous layer are shown during the initial (1, dashed lines), intermediate (2, dotted lines) and final (3, dash-dot lines) stages of indentation, and during (4, dashed lines in b) and after (5, dotted lines in b) unloading. In the final configuration (5), note the adherent amorphous layer on the crystalline indenter.

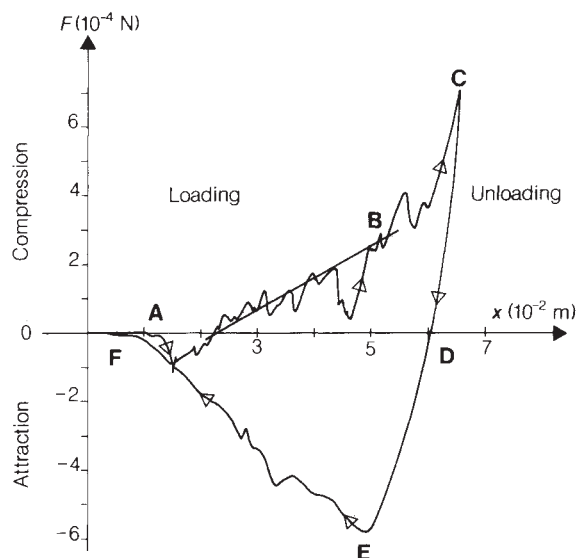


Fig. 3 Loading and unloading curves for indentation experiment. Loading (A-C): after an initial attractive stage governed by adhesion (A), the force F increases; first gradually, in proportion to displacement x (A-B, governed by amorphous layer behaviour), and then more steeply (B-C, governed by substrate behaviour). Unloading (C-F): elastic recovery (C-D) is followed by traction (D-E), and then by contraction of the amorphous layer (E-F).

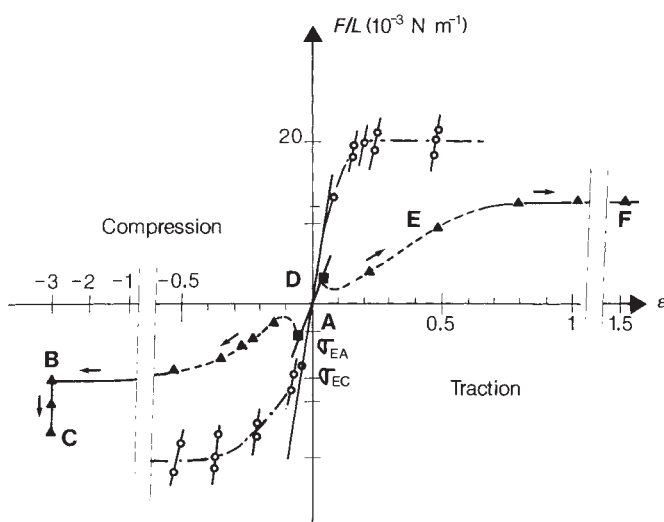


Fig. 4 Mechanical behaviour of crystalline (○) and amorphous (▲) rafts. Data points lying on solid lines were obtained from the analysis of the indentation process, those on broken lines from additional traction and compression experiments. σ_{EA} and σ_{EC} are the elastic limits of the amorphous and crystal tests, respectively. During the process of indentation, the mechanical behaviour of the amorphous layer follows points A, B, C, D, E, F.

two distinct stages, affecting first the amorphous layer and then the substrate.

On unloading, the first step is elastic recovery (C–D, Fig. 3) of the crystalline rafts. This is followed by traction, transmitted through the amorphous layer (D–E), which finally decreases (E–F) as the amorphous material contracts.

In order to treat the stress and strain data quantitatively we define the mean stress to be the force F divided by the specimen breadth L (Fig. 2a). We therefore measure, for each displacement x , the dimensions h_I of the indenter, h_L of the layer and h_S of the substrate, measured along the symmetry axis δ – δ' (Fig. 1). The values $\epsilon_i = (h_{i0} - h_i)/h_{i0} = X_i/h_{i0}$ ($i = I, L, S$) give the uniaxial deformation of each layer. The measurement of the breadth L of the indenter then yields the two-dimensional stress F/L .

The above observations suggest that the amorphous material behaviour accounts for the bulk of the stress $F/L = f(x_L/h_{L0})$ during most of the experiment. To investigate the behaviour of the crystalline raft, a compressive and tensile test curve for a uniform sheet of large bubbles is also plotted (Fig. 4).

We find that the elastic modulus (see also ref. 12) and yield stress of the amorphous material are less than those of the crystalline material. Instabilities are clearly shown in the stress–strain curves of the crystalline material, in accordance with previous work¹³. Finally, the hardness measurement shows that the mean pressure is proportional to the yield value, with the constant of proportionality depending on the shape of the indenter.

Received 1 December 1985; accepted 5 February 1986.

1. Tabor, D. *The Hardness of Metals* (Clarendon, Oxford, 1951).
2. Atkins, A. G. *Metal Sci. J.* **16**, 127–137 (1982).
3. Johnson, K. L. *J. Mech. Phys. Solids* **18**, 115–126 (1970).
4. Pollock, H. M., Maugis, D. & Barquins, M. *IMS/ASTM Microindentation Hardness Testing Symp.* 1984 (ASTM, in the press).
5. Tonck, A., Loubet, J. L. & Georges, J. M. ASLE Preprint no. 85 TC-2A-1, Tribology Conf., Atlanta (1985).
6. Bragg, L. & Nye, J. F. *Proc. R. Soc. A* **190**, 474–481 (1949).
7. Bragg, L. & Lomer, W. M. *Proc. R. Soc. A* **196**, 171–181 (1949).
8. Lomer, W. M. *Proc. R. Soc. A* **196**, 182–194 (1949).
9. Argon, A. S. & Shi, L. T. in *Proc. Fall Meet. Metall. Soc. AIME* (ed. Vitek, V.) 279–303 (1982).
10. Georges, J. M. & Meille, G. *Proc. JSLE int. Tribology Conf.*, Tokyo, 885–890 (1985).
11. Maeda, K. & Takeuchi, S. *Phil. Mag.* **A44**, 643–656 (1981).
12. Simpson, A. W. & Hodgkinson, P. *Nature* **237**, 320–322 (1972).
13. Maeda, K. & Takeuchi, S. *Phys. Status Solidi* **A49**, 685–696 (1978).

A novel, non-halogen-based chemical oscillator

Árpád Nagy & Ludovít Treindl

Department of Physical Chemistry, Comenius University,
842 15 Bratislava, Czechoslovakia

Although a kinetic bistability in the reduction of permanganate by hydrogen peroxide¹ as well as in the reduction of permanganate by oxalate^{2,3} in a continuous-flow tank reactor (CTR) has recently been observed, a permanganate chemical oscillator has not yet been described. So far, the only oscillator known has been a bromate one containing permanganate, Mn^{2+} ions and oxalate as a substrate⁴. The discovery of new chemical oscillators which are essentially different from those described previously is of great importance, as it supplies further information concerning the exact mechanistic requirements for chemical oscillations to occur. We report here the conditions under which a permanganate chemical oscillator functions in a CTR.

Our novel chemical oscillator consists of an aqueous solution of potassium permanganate, hydrogen peroxide and phosphoric acid. During the reaction in a CTR, the platinum redox-electrode potential and polarographic limiting current oscillate. A periodic change of colour of the bulk of the solution from violet to pink during the oscillation is also observed. Figure 1 shows the typical oscillatory pattern of the Pt redox-electrode potential and of the polarographic limiting current.

We used a glass CTR (40 ml volume) with a mantle connected to a TB 150 ultrathermostat (Medingen), and monitored the oscillatory reaction using a Pt redox electrode (Radelkis, Hungary), or polarographically by means of a static Pt electrode in flowing, well-stirred solution. We applied a potential of 0.0 or +1.0 V to a separated-mercurousulphate (1.0 M) electrode. The polarographic current at a potential of 0.0 V results mostly from the reduction of MnO_4^- ions; at +1.0 V it results from the anodic oxidation of Mn^{2+} ions. We were able to monitor the chemical oscillations at both potentials. The experiments were done with a stirring frequency of 400 min^{-1} . The flow of solution through the CTR was driven by a peristaltic pump, type PP 2-15 (ZALIMP, Poland). The potentiometric and polarographic measurements were made using a precision digital pH meter, OP-208/1, and a polarograph, type OH 102 (Radelkis, Hungary), respectively.

Our system (which comprises aqueous permanganate, hydrogen peroxide and phosphoric acid) exhibits a kinetic bistability either between two stationary states or between an oscillatory and a stationary state (Fig. 2). Following a stationary polarographic limiting current or a Pt redox-electrode potential, a kinetic bistability between two stationary states can be observed (Fig. 2a), if both reactants are in a stoichiometric ratio or if there is excess hydrogen peroxide. This kinetic bistability is similar to that of the permanganate–oxalate system², which we were able to reproduce in our CTR. Following the conventions of De Kepper, Epstein and Kustin^{1,5}, we use the terms 'thermodynamic branch' and 'flow branch' to describe the permanganate–hydrogen peroxide bistability. The set of states with low polarographic current (PC) is referred to as the reduced steady state (RSS) because manganese is mainly in its reduced form; this set of states corresponds to the thermodynamic branch. The set of states with high PC is referred to as the oxidized steady state (OSS), as manganese is in its oxidized form, $Mn(VII)$. This set of states corresponds to the flow branch. The two branches overlap over a large range of k_0 values. The arrows in Fig. 2 indicate the transitions from one steady state to another.

Figure 2b shows the behaviour of the system when potassium permanganate is in a stoichiometric excess. Starting from the

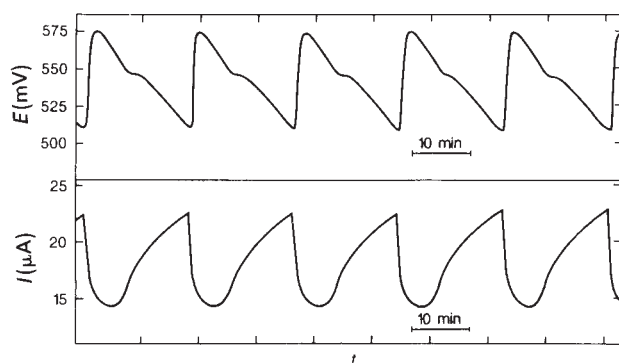


Fig. 1 Chemical oscillations in the permanganate-hydrogen peroxide system (2×10^{-4} M KMnO_4 , 3×10^{-4} M H_2O_2 , 10^{-3} M H_3PO_4 , $k_0 = 1.8 \times 10^{-3} \text{ s}^{-1}$, temperature 10°C). The polarographic current was measured with a potential of 0.0 V applied to the 1.0 M mercurousulphate electrode.

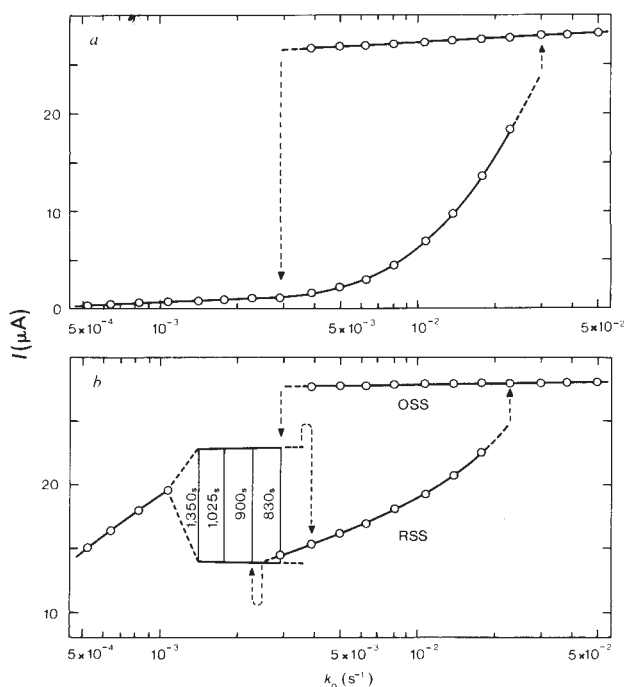
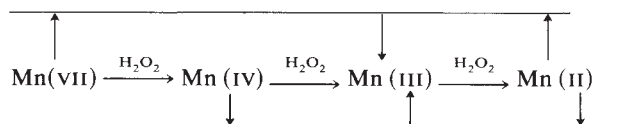


Fig. 2 Plots of polarographic current (I) versus flow rate (k_0) for the KMnO_4 - H_2O_2 system. In *a* (2×10^{-4} M KMnO_4 , 6×10^{-4} M H_2O_2 , 10^{-3} M H_3PO_4 , temperature 10°C , $E = 0.0$ V) the system displays kinetic/bistability; in *b* (2×10^{-4} M KMnO_4 , 3×10^{-4} M H_2O_2 , 10^{-3} M H_3PO_4 , temperature 10°C , $E = 0.0$ V) the system has two steady states and an oscillatory state. The system behaviour is described in detail in the text.

highest value of k_0 , the system is in an oxidized steady state. On decreasing k_0 to $3 \times 10^{-3} \text{ s}^{-1}$, the system passes from an OSS to an oscillatory state, which persists down to $k_0 = 1.4 \times 10^{-3} \text{ s}^{-1}$. On the other hand, on increasing k_0 from $1.4 \times 10^{-3} \text{ s}^{-1}$ to $3 \times 10^{-3} \text{ s}^{-1}$, the system passes from oscillation to a lower set of steady states. The system does not oscillate at $k_0 = 3 \times 10^{-3} \text{ s}^{-1}$ if returning from the lower set of steady states.

In the conditions described here, the system exhibited sustained oscillations even when the inner wall of the CTR was covered with a film of paraffin. Also, we were able to follow the periodic colour changes of the bulk of the solution even when there was no Pt electrode immersed in the solution. Thus, neither a glass surface nor the presence of platinum was necessary for the chemical oscillations to occur.

The oscillations observed here may arise from the following feedback mechanism:



Further studies are under way in our laboratory on the detailed mechanism of this chemical oscillator, focusing on the role of Mn intermediates and phosphoric acid. We believe we have found the first permanganate chemical oscillator, which might herald the beginning of an era of chemical oscillators based on transition-metal chemistry.

Received 26 November 1985; accepted 14 January 1986.

1. De Kepper, P., Ouyang, Q. & Dulos, E. in *Non-Equilibrium Dynamics in Chemical Systems* (eds Vidal, C. & Pacault, A.) 44-49 (Springer, Berlin, 1984).
2. Reckley, J. S. & Showalter, K. *J. Am. chem. Soc.* **103**, 7012-7013 (1981).
3. Epstein, I. R., Dateo, C. E., De Kepper, P., Kustin, K. & Orbán, M. in *Nonlinear Phenomena in Chemical Systems* (eds Vidal, C. & Pacault, A.) 188-191 (Springer, Berlin, 1981).
4. Dallison, A. M., Macer, D. R. J. & Rodley, G. A. *Inorg. chim. Acta* **76**, L219-L221 (1983).
5. De Kepper, P., Epstein, I. R. & Kustin, K. *J. Am. chem. Soc.* **103**, 6121-6127 (1981).

Intensity modulation in SAR images of internal waves

D. R. Thompson & R. F. Gasparovic

Johns Hopkins University, Applied Physics Laboratory, Laurel, Maryland 20707, USA

During the SAR Signature Experiment¹ (SARSEX), conducted in 1984 in the New York Bight off the eastern coast of the United States, synthetic aperture radar (SAR) images of the ocean surface and concurrent ground-truth measurements were collected in order to test quantitatively various imaging theories. These theories²⁻⁴ predict that the intensity modulations observed in SAR images of internal waves are produced by variations in the small-scale surface roughness induced by spatial variations in the internal-wave surface current field. Analysis of the SARSEX data indicates that the imaging theories can explain the observed modulation for L-band (~ 24 -cm) SAR wavelengths, but under-predict by almost an order of magnitude the observed X-band (~ 3.2 -cm) modulation. To explain this latter discrepancy, we hypothesize a two-step mechanism in which the observed X-band modulation results from additional small-scale roughness produced by the strong perturbation of the metre-scale surface waves by the internal wave current field.

The SARSEX data¹ include L- and X-band SAR images of the ocean surface from various incidence angles and viewing directions, and for the first time, concurrent ground-truth measurements of the wind velocity, surface current velocity and vertical temperature profile for each image. Figure 1 shows typical L- and X-band images of the surface manifestation of a group of internal waves propagating towards the top (range direction) of the image. The SAR incidence angles to these waves range from 25° to 47° which correspond to Bragg wavenumbers k_B (equal to twice the projection of the radar wavenumber on the horizontal surface) of 22 - 38 rad m^{-1} for L-band and 166 - 287 rad m^{-1} for X-band. The research vessels *Bartlett* and *Cape*, seen near the bottom of the images, collected ground-truth data by traversing the waves on parallel tracks towards the top of the scene. Note that both bright and dark bands are associated with each internal wave in the L-band image, but that only bright bands appear at X-band. An acceptable SAR imaging theory, when coupled with the pertinent environmental data, should be able to explain quantitatively the features observed in these images.

For the radar frequencies and incidence angles appropriate to Fig. 1, the radar backscatter from the ocean surface is given by the Bragg scattering limit^{5,6}. In this limit, the SAR intensity modulation (or relative intensity) for range-travelling waves is given by

$$M(k_B, x) = \frac{S(k_B, x) + S(-k_B, x)}{S_{eq}(k_B) + S_{eq}(-k_B)} \quad (1)$$

where the surface-wave height spectral density $S(k_B, x)$ is a function of position (x) due to the presence of the surface current field. The quantity $S_{eq}(k)$ is the equilibrium (no-current) spectrum. The problem is to understand how $S_{eq}(k)$ is modulated by the surface currents.

As pointed out by Hughes⁴ and Alpers², the time and spatial variations of the internal-wave current field are much slower than those of the relevant surface waves, so that a Wentzel, Kramer, Brillouin-type theory⁷ may be used to describe their interaction. For this study we adopt an action-balance equation proposed by Hughes⁴ of the form

$$\frac{dN(k; x, t)}{dt} = \beta(k) N(k; x, t) \left[1 - \frac{N(k; x, t)}{N_{eq}(k)} \right] \quad (2)$$

where the action spectrum $N(k; x, t) = \rho[\omega_0/|k|]S(k; x, t)$, ρ is the fluid density, ω_0 is the intrinsic frequency of a wave in a local rest frame given by $\omega_0^2 = gk[1 + (k/363)^2]$, with g being the gravitational acceleration, and the $(k/363)^2$ term accounts for surface tension effects. The time dependence of k and x in equation (2) is given by the coupled system

$$\begin{aligned} \frac{dx_i}{dt} &= \frac{\partial \omega_0}{\partial k_i} + U_i \\ \frac{dk_i}{dt} &= -k_j \frac{\partial U_j}{\partial x_i} \end{aligned} \quad (3)$$

where $U_i(x, t)$ is the surface current field, the indices run from 1 to 2, and repeated indices are summed. This system defines trajectories in four-dimensional phase space along which equation (2) is valid. The size of the source term on the right-hand side of equation (2) depends on the function $\beta(k)$, the so-called relaxation rate. It is an increasing function of k and wind speed, but there is considerable uncertainty as to its exact functional form^{4,8}. In the present study we use the expression for $\beta(k)$ given by Hughes⁴. With the relaxation time $\tau \equiv \beta^{-1}$, we find, for the conditions of the SARSEX images in Fig. 1, that τ ranges from 10 to 20 s at L-band and from 1 to 2 s at X-band.

With the definition $S(k, x, t)/S_{eq}(k) = [1 + P(k; x, t)]^{-1}$, we can integrate equation (2) to obtain $P(k; x, t)$ in the form

$$\begin{aligned} P(k; x, t) &= N_{eq}(k) \int_{-\infty}^t k'_j \frac{\partial U_j(x')}{\partial x'_i} \frac{\partial Q_{eq}(k')}{\partial k'_i} \\ &\quad \exp[-\int_{t'}^t \beta(k'') dt''] dt' \end{aligned} \quad (4)$$

where $N_{eq}(k) = \rho[\omega_0/|k|]S_{eq}(k)$, $Q_{eq} = 1/N_{eq}$, and the time dependence of the primed and double-primed variables is given by equations (3). For the results presented below, $S_{eq}(k)$ is given by a modified Pierson equilibrium spectrum⁹, with a $\cos^4[(\theta_k - \theta_w)/2]$ angular dependence where θ_w is the wind direction. No assumptions concerning the properties of the current fields have been made in obtaining equation (4). For the case where $|U| \ll \partial \omega_0 / \partial k$ and $|\partial U / \partial x| \ll \beta$, it can be shown that equation (4) yields the perturbation given by Alpers².

Experimentally determined surface currents and winds were input to the model described above, and the calculated intensity modulations were compared with those measured from the SAR images. The measured currents were parameterized in the form

$U_0 \text{sech}^2[K(x - C_1 t)]$, with typical values for the peak current U_0 , phase speed C_1 and peak gradient for the waves shown in Fig. 1 being ~ 0.3 – 0.6 m s^{-1} , 0.65 – 0.75 m s^{-1} and 0.002 – 0.008 s^{-1} , respectively. The wind speed was 6 m s^{-1} towards 145° with respect to the direction of internal wave propagation.

Figure 2 compares the measured L- and X-band intensity modulations along the track of *Bartlett* over waves 2, 3 and 4 (solid curves) with the predictions from the wave-current interaction model (dashed curves). The measured relative intensities were obtained from the digital SAR imagery by averaging over about 145 m in azimuth, smoothing over 30 m in range and normalizing so that the mean background intensity is unity. As Fig. 2a shows, the general agreement between measurement and prediction at L-band is quite encouraging, although the maximum measured modulation is somewhat under-predicted. Note that the calculated modulation reproduces both the enhancement and reduction in intensity observed in the L-band images. In fact, the computed modulation is roughly proportional to the horizontal current gradient, in agreement with Alpers' theory². Thus, we feel that, within the uncertainties inherent in the ground-truth measurements and the exact value of the relaxation rate, the wave-current interaction theory outlined above can adequately explain the L-band SAR imaging of internal waves.

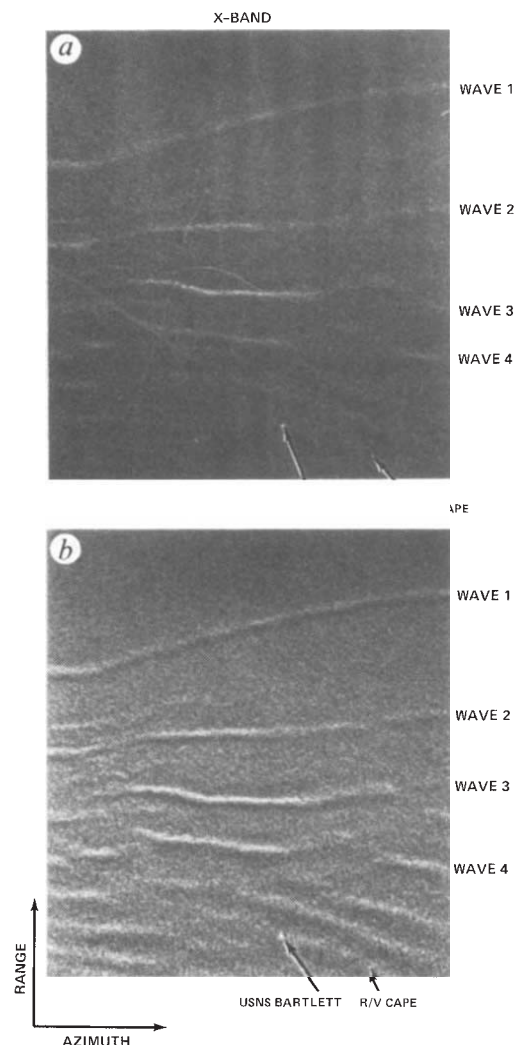


Fig. 1 Concurrent X-band (a) and L-band (b) SAR images of internal waves recorded during the SARSEX.

The situation at X-band is not so clear, however (Fig. 2b). The measured X-band relative intensity maxima have nearly the same values and occur at nearly the same position as the corresponding L-band maxima, but the pronounced regions of reduced intensity observed at L-band are not so apparent. The computed X-band modulations, on the other hand, differ from the background by only a few per cent. This is because the source term on the right-hand side of equation (2) is more effective in forcing the X-band spectral components towards equilibrium. This under-prediction cannot be removed by simply

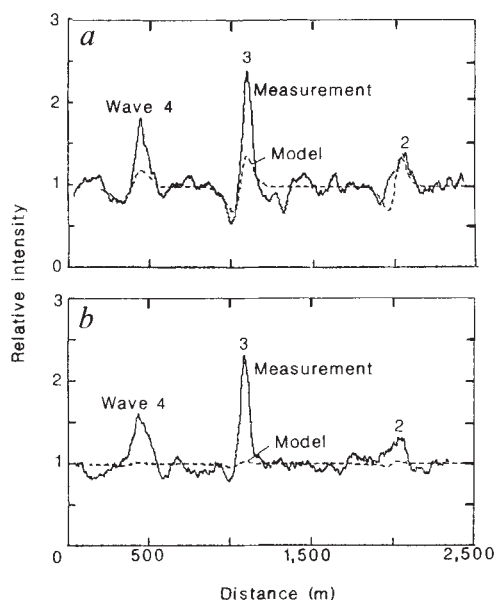


Fig. 2 Comparison of the measured (solid curves) and computed (dashed curves) L- and X-band relative intensity for waves 2-4 as discussed in the text. *a*, L-band; *b*, X-band.

adjusting the X-band relaxation rate to a value similar to that at L-band without sacrificing the intuitive feeling that shorter surface waves should respond more quickly to wind forcing than should longer ones. Furthermore, such an adjustment of the X-band relaxation rate would produce significant regions of decreased intensity which are not apparent in the X-band images.

These observations indicate that although the presence of the internal-wave surface current does indeed produce enhanced roughening of the surface at X-band scales, this roughening is not due to direct modulation of X-band Bragg waves as described by equation (2). Since it is well known that steep metre-scale gravity waves tend to lose energy to smaller scales through nonlinear interactions and ultimately through breaking¹⁰, we propose that the observed X-band modulation is the result of a two-step process in which small-scale roughening is generated due to the perturbation of the longer wind waves by the internal-wave surface current field. Hughes and Gower³ have alluded to a similar hypothesis as a possible explanation for observed differences in L- and X-band SAR images of internal waves in the Georgia Strait.

For this hypothesis to be credible, the longer wind waves must be substantially modulated by the internal-wave surface current at roughly the same position as are the L-band Bragg waves. We have investigated this point by computing the perturbation of a band of metre-scale waves by the surface current due to internal wave 3. Results are presented in Fig. 3c for wavelengths ranging from 0.21 m (L-band Bragg wave) to 2π m; the family of curves between $\lambda = 2\pi$ and $\lambda = 0.2\pi$ m corresponds to wavenumbers between 1 rad m⁻¹ and 10 rad m⁻¹ in steps of 1 rad m⁻¹. This plot shows that the position of maximum perturbation for the longest waves is near the peak of the internal-wave surface current (Fig. 3a). As the wavelength decreases, the position of the maximum shifts towards the position of the maximum (negative) gradient (Fig. 3b). Moreover, peak modulations of two to three times the equilibrium value occur throughout this region of the internal wave phase. Although we cannot quantify the degree to which the small-scale surface roughness is affected by this modulation of the longer metre-scale waves, it seems likely that it is large enough to cause a significant increase, for example, in the density of parasitic capillary waves or small-scale, locally breaking waves, thereby affecting the X-band SAR return. The region of large modulation occurs at roughly the same location as the maximum L-band perturbation, consistent with the observation from the SAR data. The modulation of the longer waves might also produce additional L-band scatterers which could explain the slight under-prediction of the maxima in the L-band SAR signatures discussed earlier. Finally, we see from Fig. 3 that there is an extended region where the metre-scale waves show a modulation less than unity. In this region, there should be no effect on the small-scale roughness from the two-step process discussed above, and the magnitudes of the X-band Bragg components should therefore be comparable to their equilibrium values. This would explain why no noticeable dark bands, such as those observed at L-band, are visible on the X-band images.

We hope to test these ideas further by analysing additional SAR imagery under environmental and imaging conditions producing a different character for the longer-wave modulations. Such analyses should add significantly to our understanding of the physics governing SAR imaging of ocean currents, and perhaps eventually lead to the development of a tool for general internal wave monitoring.

We thank E. Kasischke (Environmental Research Institute of Michigan) for the SAR data and for discussions; also our colleagues at Johns Hopkins/APL including J. Apel, B. Gotwols and A. Brandt for useful discussions, J. Kerr and B. Raff for processing the SAR images, and N. Curtiss and K. McMakin for preparation of the manuscript. The work was sponsored by the ONR.

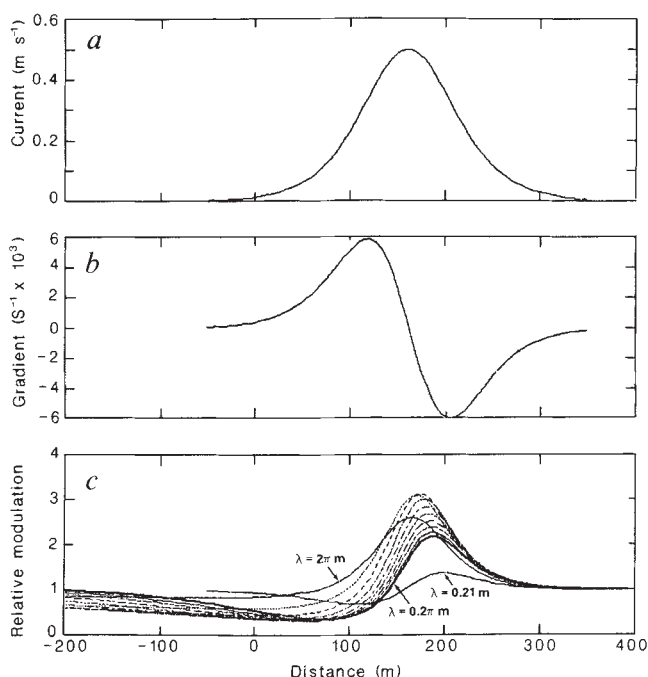


Fig. 3 Comparison of the spatial behaviour of the current and current gradient determined for wave 3 with that of the computed spectral modulation at several different wavelengths. *a*, Surface current; *b*, surface current gradient; *c*, spectral modulation.

Received 7 August 1985; accepted 15 January 1986.

1. Gaspavovic, R. F. & Apel, J. R. *Proc. 1985 Int. Geoscience and Remote Sensing Symp.*, IEEE Catalog 85ChZ162-6, 1123-1128 (1985).
2. Alpers, W. *Nature* **314**, 245-247 (1985).
3. Hughes, B. A. & Gower, J. F. R. *J. geophys. Res.* **88**, 1809-1824 (1983).
4. Hughes, B. A. *J. geophys. Res.* **83**, 455-465 (1978).
5. Valenzuela, G. R. *Radio Sci.* **3**, 1057-1066 (1968).
6. Wright, J. W. *IEEE Trans. Antennas Propagation*, AP-14, 749-754 (1966).
7. LeBlond, P. H. & Mysak, L. A. *Waves in the Ocean* (Elsevier, New York, 1978).
8. Plant, W. J. *J. geophys. Res.* **87**, 1961-1967 (1982).
9. Bjerkas, A. W. & Riedel, F. W. *Tech. Mem. TG-1328* (Johns Hopkins University Applied Physics Laboratory, 1979).
10. Phillips, O. M. *The Dynamics of the Upper Ocean* (Cambridge University Press, 1977).

Direct indication of pore-water advection from pore pressure measurements in Madeira Abyssal Plain sediments

P. J. Schultheiss & S. D. McPhail

Institute of Oceanographic Sciences, Brook Road, Wormley, Godalming GU8 5UB, UK

The difference in pressure between the sea floor and a point in the underlying sediment is an indicator of the velocity of pore-water flow. This differential pore pressure has now been measured at five sites using a free-fall piezometer (pop-up pore pressure instrument, PUPPI) in a distal turbidite province of the Madeira Abyssal Plain. We found that at one of the five sites the pressure was slightly positive, at two others it was effectively zero, but at the remaining two sites the pore pressure was significantly negative (-450 and -120 Pa at 4 m depth). The implication is that at these last two sites water is moving downwards at rates of 3.1 and 0.8 mm yr^{-1} (calculated using laboratory-measured hydraulic conductivities from core samples). If this downward movement is part of a hydrogeological system which includes upward movement of pore water elsewhere, there are obvious implications for schemes to dispose of highly radioactive waste by burial in deep-sea sediments.

Pore pressure measurements in soils and sediments provide an essential component in the evaluation of the effective stress condition which governs the material behaviour. The effective stress is given by $\sigma' = \sigma - u$, where σ is the total stress and u is the pore pressure. Thus pore pressure measurements in ocean sediments are of interest to those concerned with offshore construction. In addition, many geological processes (for example, consolidation, lithification, sediment stability and liquefaction potential) will be influenced by the state of effective stress and hence by the pore pressure. Another process which will be dominated by pore pressure gradients is the mass movement of pore water through the sediments. The PUPPI was designed to investigate the possible advection of pore waters at sites designated as study areas for the feasibility of radioactive waste disposal. Pore-water advection is one of many factors that could potentially accelerate or retard the release of dissolved radionuclides to ocean waters from buried canisters.

It is generally accepted that the anomalously low values of conductive heat flux observed in areas close to active spreading centres can be accounted for by hydrothermal circulation through the highly permeable ocean crust into the sea; this process has been observed to occur in young oceanic crust¹⁻⁴. Evidence for the existence of pore-water advection in deep-sea sediments comes mainly from nonlinear temperature profiles measured at several locations⁵⁻⁹. Vertical pore-water advection velocities calculated from such temperature profiles and from pore-water chemistry have been compared at three sites in the equatorial Pacific¹⁰. It was found that the velocities calculated using sulphate diffusion-convection-reaction models were two to three orders of magnitude lower than velocities calculated from the nonlinear temperature profiles, suggesting that mechan-

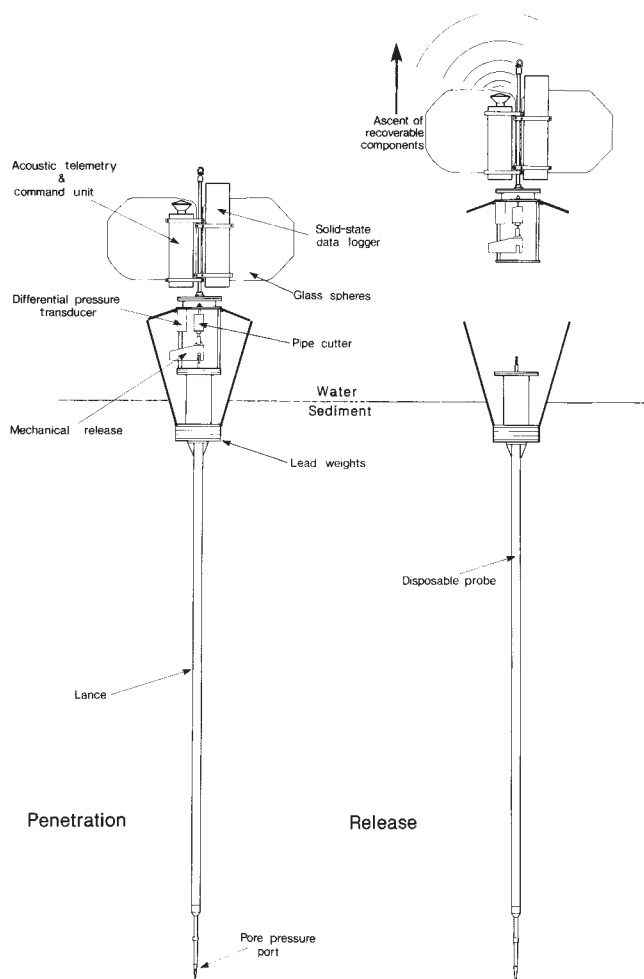


Fig. 1 Schematic diagrams of the pop-up pore pressure instrument (PUPPI): left, after its free-fall through the water column and penetration into the sediments; right, after release from the sea floor, showing recoverable and disposable components.

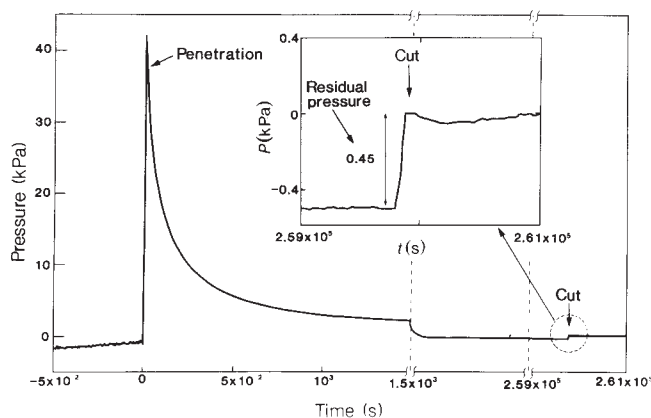
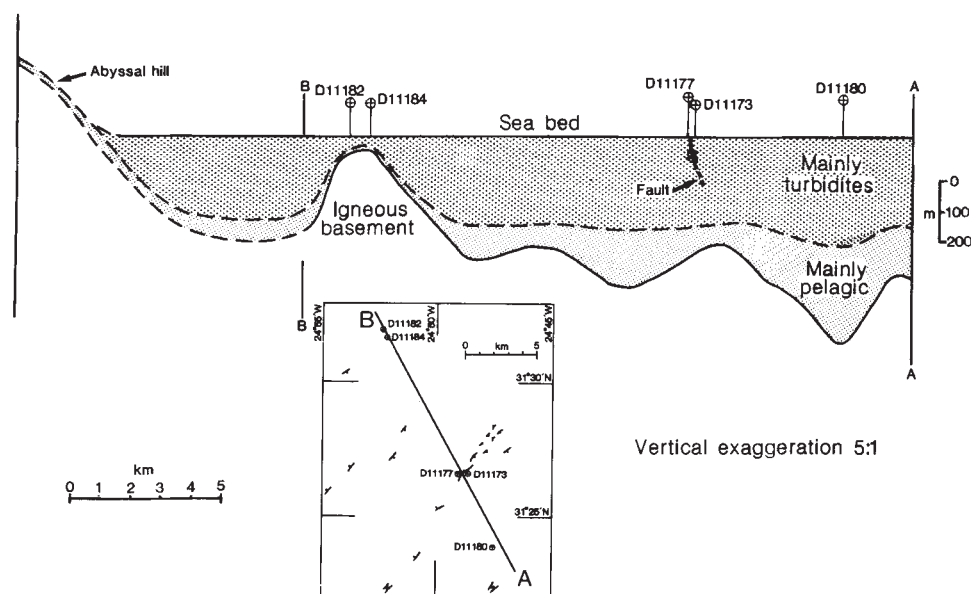


Fig. 2 Pressure-plot for PUPPI station D11184 showing the insertion pressure pulse at zero time (penetration) followed by the decay curve and the cut. Note the changes in the timescale. The inset shows an expanded record around the cut revealing a -0.45 -kPa residual differential pressure.

isms other than pore-water advection are responsible for the profiles¹¹. In addition it is unlikely, with currently available techniques, that a resolution of better than 100 mm yr^{-1} could be obtained using temperature profiles. Studies of pore-water advection relevant to the disposal of high-level radioactive waste must have a resolution of better than 1 mm yr^{-1} .

Fig. 3 Section through the Great Meteor East study area showing the location of the five PUPPI sites in relation to faults, basement topography and sediment thickness. Inset shows location of section A-B and station positions.



The movement of water through a permeable medium is described by Darcy's law: $Q = KA\Delta u / \eta L$, where Q is the volume flow-rate of a liquid of viscosity η caused by a pressure difference Δu over a length L in a cross-section A of a porous medium with specific permeability K . When water is the permeant it is customary to define the hydraulic conductivity: $k = K\gamma g / \eta$, where γ is the density of the fluid and g the acceleration due to gravity. The velocity of an element of water in a permeable medium is $v = Q/A$; therefore, $v = k\Delta u / \eta \gamma L$, where n is the fractional porosity of the sediment. Values of k and n can be measured from core samples in the laboratory but the differential pore pressure must be measured *in situ* by a piezometer.

The PUPPI (Fig. 1) was designed at the Institute of Oceanographic Sciences¹² to measure residual pore pressures at a depth of 4 m in soft sediments with a final resolution of 10 Pa. It is a free-fall instrument that can be ballasted to penetrate a range of sediment types in water depths of up to 6,000 m. In its present configuration it has a single pressure port near the end of the lance, connected to a differential pressure transducer (the other port is open to the sea). It therefore measures the pore pressure with respect to sea-bed hydrostatic pressure. A programmable solid-state data logger is used to record the pressure data at a maximum rate of 0.5 sample per s. An accelerometer record of the penetration event enables the depth of penetration to be calculated. Recovery is accomplished using an acoustic command that activates a release mechanism which simultaneously cuts the pipe connecting the pressure port to the transducer. The lance, ballast and cone assembly are left on the sea floor; the buoyant instrument package ascends to the surface for recovery.

A sample pressure-time record is shown in Fig. 2 (note the changes in timescale which are necessary to illustrate the behaviour). When the instrument penetrates the sea floor, a pore-pressure transient (~ 40 kPa) results from the large radial stress imposed on the sediment by lance insertion and from the small vertical stress due to the submerged weight of the PUPPI (~ 75 kg). This pressure transient decays with time as the sediment undergoes primary consolidation. After 10 min it has decayed to 10%, and after 3 h less than 1%, of the peak value. No detectable decay occurs during the final 24 h; hence it is assumed that the pore pressure caused by the instrument has decayed to zero.

To obtain pore pressure measurements of the required accuracy, the transducer must be accurately calibrated. Laboratory experiments have shown that changes in ambient temperature and pressure cause only a small ($<1\%$) change in the

total range; however, for maximum accuracy an *in situ* zero-calibration method is used. Pressure data are logged rapidly near the time when the pore pressure pipe connecting the transducer to the pressure port is cut. The reading after the cut must be zero and any change in reading near the time of the cut, after averaging to remove the signal from tidal cycles, indicates the existence of a differential pore pressure in the sediments. The data shown in Fig. 2 are plotted with reference to the zero pressure reading obtained after the cut. During the 500 s before penetration, the zero point on the transducer drifted by ~ 1 kPa as a result of temperature and pressure changes. It drifted a further 0.5 kPa before the pipe was cut 2 days later.

Data were collected on *Discovery* cruise 153 from five sites within an area known as the GME (Great Meteor East) study area. Figure 3 shows a section through the area, based on the interpretation of coring and seismic reflection profiling data¹³. The inset shows the station positions. The area is an abyssal plain, comprising ~ 90 -Myr-old basement overlain by a sequence of pelagic sediments, which are in turn overlain by a series of thick distal turbidites¹⁴. Within the turbidite sequence there are a large number of faults which are clearly revealed on 3.5-kHz reflection profiles¹³. The PUPPI stations cover three major areas: (1) D11180 is in an area of thick sedimentary cover (~ 700 m); (2) D11173 and D11177 lie within 50–200 m of the fault line (determined with bottom transponders for navigation); (3) D11182 and D11184 are in an area of thin sedimentary cover (~ 75 and 50 m) over a sub-bottom basement high that does not outcrop. The results are summarized in Table 1. Two sites show significant negative residual pressure gradients (D11177 and D11184). As shown in Fig. 3, D11177 is ~ 50 m from a fault, on the downthrown side, whereas D11173 is ~ 200 m from the same fault on the upthrown side. D11184 lies very close to the peak of the sub-bottom basement high, where there is ~ 50 m of sediment cover, whereas D11182 lies on the flank, where the cover thickness is ~ 75 m. One site (D11180) shows a small positive residual pressure, but the instrumental precision is not sufficient to distinguish this result from a differential pressure of zero.

The advection velocities calculated for the two sites which show negative residual pressures (D11177 and D11184) are -0.8 mm yr^{-1} and -3.1 mm yr^{-1} respectively. The negative sign indicates that the movement is downward. These velocities are calculated using laboratory-measured values for the hydraulic conductivity (corrected for *in situ* conditions using density and viscosity data¹⁵) and porosity (see Table 1).

In these experiments the water which was used to fill the pipe before deployment had a measured density which could exceed

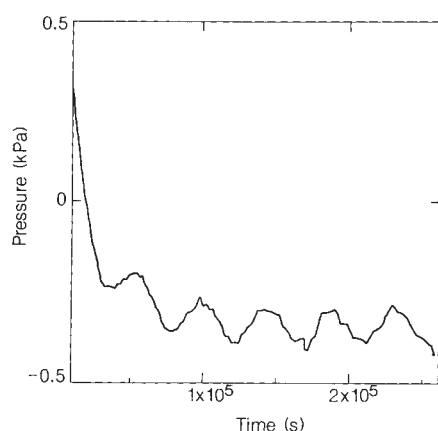


Fig. 4 Expanded pressure-time plot for PUPPI station D11184, showing the attenuated tidal cycles recorded as a result of the finite compliance of the differential pressure transducer.

that of the bottom water by as much as 0.18%. This density difference could cause an apparent residual pressure of -78 Pa, which translates into an apparent advection velocity of -0.5 mm yr^{-1} . The advection velocities listed in Table 1 might therefore have to be increased by up to 0.5 mm yr^{-1} ; this would not change the overall interpretation of the data.

The largest source of potential error is in the laboratory hydraulic conductivity measurement, and arises not from the technique used but from the possibility of unrepresentative sampling (missing fractures or open burrows) and/or disturbance¹⁶ (closing of open burrows). If this is a source of error, it is most likely that the *in situ* value will be higher than the laboratory measurement, possibly by up to an order of magnitude. If this is the case then the advection velocities at all sites could be much higher. To resolve this uncertainty, we are investigating the accurate *in situ* evaluation of hydraulic permeability. The most promising method involves the observation of an effect caused by tidal variation. Pore pressure measurements suffer from a time lag in their measurement because of the finite compliance of the transducer and the low permeability of the sediments. As a result of this, a tidal cycle appears on the pore pressure record. It is highly attenuated and possibly phase-shifted from the real tidal hydrostatic waveform. Figure 4 shows the tidal cycles observed at station D11184, which have an amplitude of ~ 120 Pa, compared with the tidal amplitude of ~ 7 kPa. The oscillations are only just within the resolution of the instrument but this could be improved by increasing the transducer resolution. A comparison with the actual hydrostatic pressure changes (monitored using a tide gauge) should enable accurate *in situ* values of hydraulic conductivity to be obtained.

The limited hydrogeological data presented for the transect shown in Fig. 3 show a downward flow of water over a sub-bottom basement high and near a fault. The questions that must be addressed for the radioactive waste feasibility study concern

the nature of the flow. Is it part of a thermally driven convection cell, and if so what is the scale of the cell and where are the discharge areas? Or is it caused by a consumption of water within the deeper sediments or basement as part of a diagenetic chemical process? Possible discharge areas are from local abyssal hills or over large areas within the thicker sedimentary sequence, but the discharge rates must be below the present resolution of the PUPPI. The positive advection velocity seen at site D11180 may be indicative of such discharge.

We thank B. Hart, A. Packwood, R. B. Whitmarsh and R. E. Kirk for significant contributions and advice during the concept and design phases of the PUPPI, and R. B. Whitmarsh, R. C. Searle, T. J. G. Francis, P. P. E. Weaver, D. Abbott, A. F. Richards and E. James for discussion and helpful criticism of the text. This research has been carried out under contract for the Department of the Environment, as part of its radioactive waste management programme.

Received 8 May 1985; accepted 23 January 1986.

1. Corliss, J. B. *et al.* *Science* **203**, 1073–1083 (1979).
2. RISE Project Group *Science* **207**, 1421–1433 (1980).
3. Macdonald, K. C. & Luyendyk, B. B. *Earth planet. Sci. Lett.* **48**, 1–7 (1981).
4. Anderson, R. N. *et al.* *Nature* **300**, 589–594 (1982).
5. Anderson, R. N., Hobart, M. A. & Langseth, M. G. *Science* **204**, 828–832 (1979).
6. Langseth, M. G. & Herman, B. M. *J. geophys. Res.* **86**, 10805–10819 (1981).
7. Becker, K. & von Herzen, R. P. *J. geophys. Res.* **88**, 1057–1066 (1983).
8. Williams, D. L. *et al.* *J. geophys. Res.* **84**, 7467–7484 (1979).
9. Abbott, D. H., Menke, W. & Morin, R. *J. geophys. Res.* **88**, 1075–1093 (1983).
10. Crowe, J. & McDuff, R. E. *Eos* **60**, 863 (1979).
11. Noel, M. J. *Geophys. J. R. astr. Soc.* **76**, 673–690 (1984).
12. Schultheiss, P. J., McPhail, S. D., Packwood, A. R. & Hart, B. *Inst. oceanogr. Sci. Rep. no.* 202 (1985).
13. Searle, R. C. *et al.* *Inst. oceanogr. Sci. Rep. no.* 193 (1985).
14. Weaver, P. P. E. & Kuijpers, A. *Nature* **306**, 360–363 (1983).
15. Schultheiss, P. J. & Gunn, D. E. *Inst. oceanogr. Sci. Rep. no.* 201 (1985).
16. Weaver, P. P. E. & Schultheiss, P. J. *Nature* **301**, 329–331 (1983).

Direct dating of the oxygen-isotope record of the last deglaciation by ^{14}C accelerator mass spectrometry

Jean-Claude Duplessy*, Maurice Arnold*, Pierre Maurice*, Edouard Bard*, Josette Duprat† & Jean Moyes†

* Centre des Faibles Radioactivités, Laboratoire mixte CNRS-CEA, F-91190 Gif-sur-Yvette, France

† Département de Géologie et Océanographie, Université de Bordeaux 1, Avenue des Facultés, F-33405 Talence, France

Major trends of Quaternary global climate are reflected in the continental ice volume changes which have been reconstructed by oxygen-isotope analysis^{1,2}. $\delta^{18}\text{O}$ records from deep-sea sediments show that the net glacial build-up occurs relatively slowly³, but that the end of an ice age occurs quickly, in less than 10,000 yr, implying a nonlinear response^{4,5} to simple Milankovitch theory⁶. The latter observation suggests that the cause of the most recent deglaciation was the maximum in summer calorific radiation at the upper latitudes of the Northern Hemisphere centred around 11,000 yr ago, a view supported by early studies⁷. Later work has produced conflicting dates, the main source of confusion being problems with obtaining accurate and reliable dates. Here, by using accelerator mass spectrometry, we have measured ^{14}C for various species of foraminifera to produce a reliable timescale for the oxygen-isotope record. Our results show that, at the end of the last ice age, continental ice sheets began to melt more than 4,000 yr before the Northern Hemisphere maximum of summer calorific radiation.

The timing of deglaciation has important consequences for the mechanisms of climate change. The earlier work suggesting

Table 1 Differential pressures and advection velocities calculated from the equation $v = k\Delta u/n\gamma gL^*$

PUPPI site no.	ΔU (Pa)	V (mm yr^{-1})
D11173	0	0
D11177	-120	-0.8
D11180	+30	+0.2
D11182	0	0
D11184	-450	-3.1

Δu , Measured residual differential pressure; v , calculated advection velocity.

* $k = 7 \times 10^{-9}$ m s^{-1} ; $n = 0.77$; $\gamma = 1,053$ kg m^{-3} ; $g = 9.81$ m s^{-2} ; $L = 4$ m.

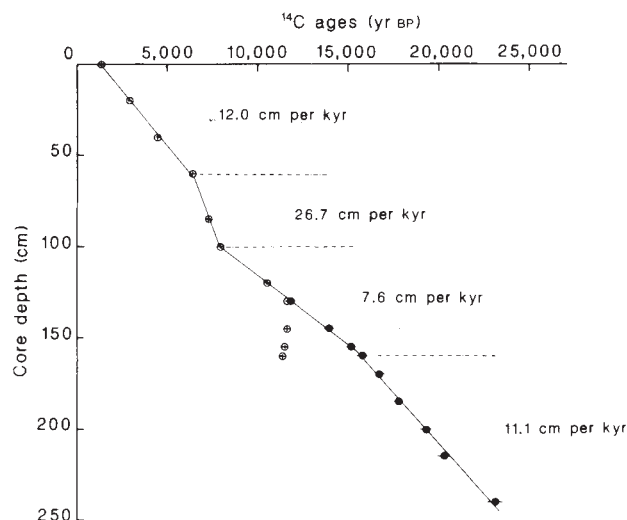


Fig. 1 AMS ^{14}C ages plotted against sample depth in core CH73-139C for two foraminiferal species, *G. bulloides* (○) and *N. pachyderma* (●).

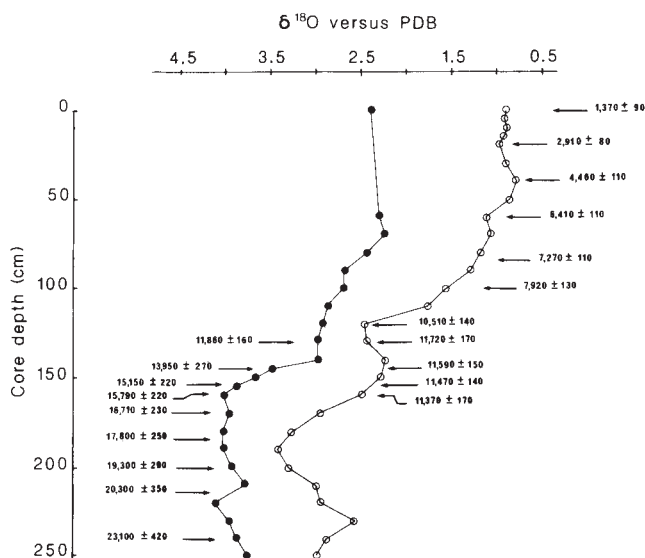


Fig. 2 Oxygen-isotope records and ^{14}C ages of the two foraminiferal species *G. bulloides* and *N. pachyderma* in the core CH73-139C.

that the maximum rate of ice-sheet disintegration was centred at $\sim 11,000$ yr ago⁷, implying that ice melting was a direct response to insolation, has been cast into doubt by more recent isotopic analysis of deep-sea sediments deposited with a high sedimentation rate in the northeastern Atlantic Ocean. This suggested that deglaciation occurred in two discrete steps, the first (termination 1a) between 16,000 and 13,000 yr ago and the second (termination 1b) between 10,000 and 8,000 yr ago⁸. This evidence implies that there was no melting of the ice sheets between 13,000 and 10,000 yr ago, the time at which fastest melting should have occurred according to earlier studies. Recent studies add even more confusion: a two-step deglaciation centred at $\sim 12,500$ yr BP was proposed by Mix⁹. Stacked oxygen-isotope data from foraminifera in tropical Atlantic sediment cores suggested that the last deglacial transition began at 14,000 yr BP^{10,11}. Other studies suggested that the deglaciation started 13,000–15,000 yr ago with one to three steps^{11–13}. There is thus a 3,000-yr uncertainty concerning the exact timing of the beginning of the last deglaciation.

The main causes of age discrepancies in deep-sea sediment cores are mainly (1) bioturbation, which mixes the upper cen-

timetres of sediment and partially blurs the palaeoclimatic record and which is particularly severe in cores with a low sedimentation rate, and (2) the presence within the North Atlantic sediment either of ice-rafted carbonate or of wind-, river- and current-transported carbonate, which are derived from pre-Pleistocene continental rocks and bias ^{14}C ages towards older values.

Accelerator mass spectrometry (AMS) provides reliable ^{14}C ages with only 1–2 mg of carbon; it thus offers a unique opportunity to obtain ^{14}C ages on pure foraminiferal samples, free of detrital contamination. We present here analyses performed on foraminiferal shells from core CH73-139C, raised from the Rockall Plateau ($54^{\circ}38' \text{ N}$, $16^{\circ}21' \text{ W}$, water depth 2,209 m). This core has a mean sedimentation rate approximately equal to 10 cm per 1,000 yr^{8,14}, which reduces (but does not cancel) the effects of bioturbation. Its isotopic and palaeoclimatic records are typical of those of the northeastern Atlantic Ocean^{8,15,16}. We determined ^{14}C ages for individual species of foraminifera using the Tandemron Accelerator in Gif-sur-Yvette¹⁷; 1,000–2,000 foraminiferal shells were examined under a binocular microscope for the two species *Neogloboquadrina pachyderma* (left coiling), which predominates during cold climatic episodes, and *Globigerina bulloides*, which is a transitional/subpolar species, abundant during interglacial conditions. The foraminiferal shells were cleaned in an ultrasonic bath, then attacked superficially with 10^{-3} M nitric acid to prevent surficial contamination from modern carbonate, and reacted under vacuum with phosphoric acid. The CO_2 evolved was converted to carbon by reaction with hydrogen in the presence of iron at 650°C (ref. 18). The $^{14}\text{C}/^{12}\text{C}$ content of the sample was compared with that of a National Bureau of Standards oxalic acid reference sample prepared from the same amount of carbon and with the same amount of iron to prevent fractionation in the accelerator source, depending on the C/Fe ratio. Other fractionations were corrected by comparing the $^{13}\text{C}/^{12}\text{C}$ ratio of the sample with that of the reference sample¹⁹. This procedure provides values for the $^{14}\text{C}/^{12}\text{C}$ ratio that are accurate to better than 1% of the modern ^{14}C content (M.A. *et al.*, in preparation). Six pre-Quaternary foraminiferal samples prepared according to this procedure had a background contamination of $0.58 \pm 0.15\%$ of the modern $^{14}\text{C}/^{12}\text{C}$ ratio, corresponding to an apparent age of 41,500 yr, which is the present limit of dating for the Gif accelerator. This background contribution was subtracted from all the measurements and represents a minor correction. Errors for the ^{14}C ages, calculated at the 1σ level, take into account the statistical error for the counts performed on both experimental and reference samples and the uncertainty on the background (M.A. *et al.*, in preparation). Finally, foraminiferal ^{14}C ages have been corrected for the apparent ^{14}C age of surface sea water—that is, 400 years.

Figure 1 plots ^{14}C ages against depth: in the upper 60 cm of the core, ^{14}C analyses of *G. bulloides* indicate a mean sedimentation rate of 12 cm per 1,000 yr. Between 60 and 100 cm depth, the beginning of the Holocene is characterized by a high sedimentation rate (26.7 cm per 1,000 yr). Below 160 cm, ^{14}C analyses of *N. pachyderma* indicate a mean sedimentation rate of 11.1 cm per 1,000 yr during glacial times before 15,100 BP. Between 130 and 160 cm, ^{14}C ages of *N. pachyderma* indicate a constant mean sedimentation rate of 7.6 cm per 1,000 yr from 12,000 to 15,000 yr BP. Between 100 and 160 cm, the ^{14}C ages of *G. bulloides* increase with core depth until 130 cm and remain constant from 130 to 160 cm. These constant age values indicate that *G. bulloides* became abundant in the northeastern Atlantic only after 11,600 yr BP and was mixed in older sediment levels by bioturbation. Only a few specimens were found below 160 cm. This abundance was sufficient for an oxygen-isotope analysis but not for ^{14}C dating. At 100, 120 and 130 cm depth, the ^{14}C ages of *G. bulloides* coincide with those derived from an extrapolation of the sedimentation rate provided by ^{14}C analyses of *N. pachyderma*. We therefore assume a constant sedimentation rate of 7.6 cm per 1,000 yr during the whole climatic transition. This

value is lower than both glacial and interglacial values, in agreement with the fall in productivity associated with the deglaciation¹⁵.

Below 130 cm, bioturbation introduces a strong age difference between the two species which increases progressively to more than 4,000 yr at 160 cm. This result clearly illustrates how a ¹⁴C analysis performed on a mixture of species may be misleading. The age of any given species analysed for stable isotope content may be drastically different from that of the bulk foraminiferal population if the species exhibit important fluctuations in abundance and are mixed with deeper layers in the sediment. AMS thus provides a new means to studying the impact of bioturbation in deep-sea sediments.

Figure 2 shows the oxygen-isotope records of *G. bulloides* and of *N. pachyderma* and the ¹⁴C ages obtained for the same species. The *G. bulloides* data show that termination 1b occurred after 10,500 yr BP, in agreement with the general recognition that this step followed the Younger Dryas cold event^{8-10,12,13}. The major development of *G. bulloides* at about 11,600 yr BP in the northeastern Atlantic Ocean coincides with the Allerod warm continental event. This result provides the first evidence that oceanic conditions may have varied during the Bolling/Allerod interstadial complex, which lasted from 13,300 to 11,000 yr BP.

The ¹⁴C ages obtained for *N. pachyderma* samples provide an accurate timescale for termination 1a. As this isotopic record is similar to, and in phase with, the benthic record of the same core within experimental errors⁸, we believe that it reflects changes of continental ice volume during the last deglaciation. Our data show that the first phase of major melting began soon after 15,800 yr BP (level 160 cm; $\delta^{18}\text{O} = 4.02$) and was complete by ~13,300 yr BP (level 140 cm; $\delta^{18}\text{O} = 2.97$). The mid-point of termination 1a is dated at ~14,000 yr BP, which suggests that the speed of deglaciation increased progressively during most of this melting episode and fell sharply to zero by its end. The pause in ice melting until termination 1b is confirmed by consistent ages measured in both of the foraminiferal species during the interval 12,000–10,500 yr BP.

By providing ¹⁴C ages for individuals of the same species as those analysed for $\delta^{18}\text{O}$, AMS allows us to minimize the effects of bioturbation and to derive the most accurate timescale for the palaeoclimatic record of deep-sea sediments. For example, the $\delta^{18}\text{O}$ values of *G. bulloides* exhibit their first decrease at a depth 20 cm below that of *N. pachyderma* in core CH73-139C. AMS ¹⁴C dates demonstrate that this 20-cm lead of the *G. bulloides* record over the *N. pachyderma* record is an artefact of bioturbation. The inconsistency between the two isotopic records disappears when both records are compared with the ¹⁴C ages measured on the same foraminiferal species (Fig. 2). In view of this remarkable age difference between different foraminiferal species in a core with an accumulation rate around 10 cm/kyr, it is not surprising that initial results from a core in the Pacific with an accumulation rate of only ~2 cm/kyr were disappointingly confusing²⁰.

The chronology of the last deglaciation that we propose here shows good agreement with the palaeoclimatic record of western Europe, which has a highly maritime climate closely linked to the temperature of the northeastern Atlantic Ocean. The internal coherence of both continental and oceanic evidence suggests that our timescale is probably exact within the uncertainties of the ¹⁴C dates, that is, ~250 yr in this time range.

Our data demonstrate that the first rapid phase of deglaciation occurred before the maximum in summer caloric insolation over the high and middle latitudes of the Northern Hemisphere at 11,000 yr BP, they support the hypothesis that rates of ice melting are not linearly linked to insolation^{5,6} and that the deglaciation resulted from the complex interaction of many feedback mechanisms. This early deglaciation suggests that among these feedback mechanisms, atmospheric CO₂ concentrations and ice dynamics may have had a major role. On one

hand, increasing concentrations of atmospheric CO₂ at the end of the glaciation^{21,22} may have contributed significantly to the forcing of the deglaciation because a comparison of $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ analyses of planktonic and benthic foraminifera from the Pacific Ocean suggests that the CO₂ increase began before continental ice began to melt²³. On the other hand, the mechanisms of ice dynamics are necessary to explain the disagreement between the two-step isotopic record suggesting a discontinuous volumetric deglaciation, and the variation of the ice-sheet areal extent between 15,000 and 7,000 yr BP which is more easily explained by a smooth progressive melting of the ice sheet²⁴. However, the seasonal disappearance of sea ice in the Norwegian–Greenland Sea soon after 18,000 yr BP²⁵ may have triggered rapid ice calving into the ocean and the down-draw of the great ice sheets²⁶.

We thank J. Labeyrie for introducing AMS at Gif-sur-Yvette; J. Antignac, E. Kaltnecker, B. Le Coat and H. Leclaire for help in the analyses; R. Lesueur, A. Castera and D. de Zertucha for their help in developing AMS; and N. J. Shackleton, L. Labeyrie and W. Dudley for useful discussions and review of the manuscript. This work was supported by CNRS (PNEDC), CEA and EEC grant CLI 085-F. CFR contribution no. 751.

Received 14 October; accepted 30 December 1985.

- Emiliani, C. *J. Geol.* **63**, 538–578 (1955).
- Shackleton, N. J. & Opdyke, N. D. *Quat. Res.* **3**, 39–55 (1973).
- CLIMAP. *Science* **191**, 1131–1137 (1976).
- Hays, J. D., Imbrie, J. & Shackleton, N. J. *Science* **194**, 1121–1132 (1976).
- Imbrie, J. & Imbrie, J. Z. *Science* **207**, 943–953 (1980).
- Milankovitch, M. R. *Serbian Acad. (Beograd) Spec. Publ.* **133** (1941).
- Broecker, W. S., Ewing, M. & Heezen, B. C. *Science* **258**, 429–448 (1960).
- Duplessy, J. C., Delibrias, G., Turon, J. L., Pujol, C. & Duprat, J. *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **35**, 121–144 (1981).
- Ruddiman, W. F. & Duplessy, J. C. *Quat. Res.* **23**, 1–17 (1985).
- Berger, W. H. *Sver. geol. Unders. Afh.* **76**(7c), 270–280 (1982).
- Mix, A. C. & Ruddiman, W. F. *Quat. Res.* (in the press).
- Berger, W. H., Killingly, J. S., Metzler, C. V. & Vincent, E. *Quat. Res.* **23**, 258–271 (1985).
- Samthein, M., Erlenkeuser, H. & Zahn, R. *Bull. Inst. geol. Basin d'Aquitaine, Bordeaux* **31**, 393–407 (1982).
- Pujol, C. thesis, Univ. Bordeaux 1 (1980).
- Ruddiman, W. F. & McIntyre, A. *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **35**, 145–214 (1981).
- Labeyrie, L. D. & Duplessy, J. C. *Palaeogeogr., Palaeoclimatol., Palaeoecol.* **50**, 217–240 (1985).
- Arnold, M., Lesueur, R., Maurice, P. & Duplessy, J. C. *3rd int. Symp. Accelerator Mass Spectrometry: AMS '84*, Zurich (1984).
- Vogel, J. S., Southon, J. R., Nelson, D. E. & Brown, T. A. *Nucl. Instrum. Meth.* **223**, 289–293 (1984).
- Beukens, R. P. & Lee, H. W. *Symp. Accelerator Mass Spectrometry, Proc. Argonne National Lab. ANL/PHY-81-1n*, 416–425 (1981).
- Andree, M. et al. *Nuclear Instruments and Methods in Physics Research* **B5**, 340–345 (1984).
- Berner, W., Stauffer, B. & Oeschger, H. *Radiocarbon* **22**, 227–235 (1980).
- Delmas, R. J., Ascencio, J.-M. & Legrand, M. *Nature* **284**, 155–157 (1980).
- Shackleton, N. J. & Pisias, N. G. in *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 303–317 (Am. Geophys. Un., Washington, DC, 1985).
- Paterson, W. S. B. *Rev. Geophys. Space Phys.* **10**, 885–917 (1972).
- Grousset, F. & Duplessy, J. C. *Mar. Geol.* **52**, M11–M17 (1983).
- Denton, G. H. & Hughes, T. *The Last Great Ice Sheets* (Wiley-Interscience, New York, 1981).

Extensive deposition of banded iron formations was possible without photosynthesis

Louis M. François

Astrophysical Institute, University of Liège, 5 Avenue de Cointe, B-4200 Cointe-Liège, Belgium

Precambrian banded iron formations (BIFs) consist of alternating layers of silica and iron minerals such as haematite, magnetite and siderite^{1,2}, but there is controversy as to the origin of the iron. According to one point of view, iron was precipitated from sea water containing Fe(II) ions in solution^{3,4}. Because the photodissociation of water vapour in the Precambrian atmosphere would have been too slow to generate enough of the most probable oxidizing agent, oxygen⁵, this process would have had to wait for the evolution of organisms producing oxygen. Cairns-Smith⁶ and

Braterman *et al.*⁷ alternatively suggested abiotic photodissociation of FeOH^+ . Through the use of a Precambrian ocean model, I demonstrate here that such a process can quantitatively account for the formation of known BIFs.

The ocean is divided into three regions (see Fig. 1): the mixed layer, the thermocline and the deep-water reservoir, and is composed of the following cations: Na^+ , K^+ , Mg^{2+} , Ca^{2+} , Fe^{2+} and H^+ ; and anions: Cl^- , SO_4^{2-} , HCO_3^- , CO_3^{2-} , OH^- and H_2BO_3^- . These ions can form various complexes. The mixed layer is 100 m thick and is at a temperature of 20 °C. Photo-oxidation of FeOH^+ occurs in this region. In the thermocline, which is 1,000 m thick, the temperature decreases linearly with depth. Total Fe(II) ($\text{Fe}^{2+} + \text{FeOH}^+ + \text{FeCl}^- + \text{FeSO}_4^0$) is transported by diffusion and upward advection (upwelling) throughout the thermocline. The deep-water reservoir is assumed to be homogeneous and has a temperature of 5 °C. Deep-water formation at high latitudes is modelled by a pipe linking the mixed layer to the deep-water reservoir and in which the water circulates downwards with a velocity equal to the advective velocity.

Following Holland³, the deep water is assumed to be saturated with respect to siderite. Langmuir⁸ reports a temperature-dependent expression for the activity solubility product. At 5 °C, the value $K_{\text{FeCO}_3} = 1.31 \times 10^{-10} \text{ mol}^2 \text{ kg}^{-2}$ is found for the concentration solubility product, using the activity coefficient of Fe^{2+} and CO_3^{2-} at an ionic strength $I = 0.7$. Fe(II) is transported upward from the deep water by advection and diffusion. In the mixed layer, Fe(II) is transformed into FeOH^+ ; this ion is photo-

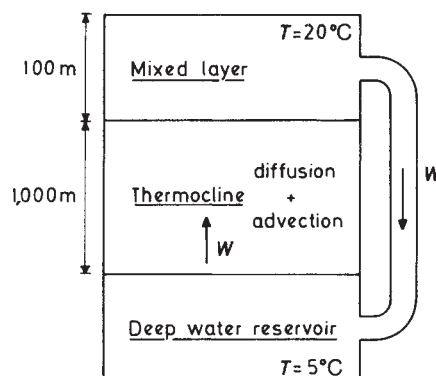


Fig. 1 Vertical structure of the ocean in the model; w is the advective velocity.

According to Braterman *et al.*⁷, the quantum yield γ of the photo-oxidation reaction lies between 0.01 and 0.05. A mean value of 0.03 is adopted. The calculated photo-oxidation rate is multiplied by 0.5 to account for the day-night alternation.

The photochemical model requires the FeOH^+ concentration in the mixed layer as an input. It may be obtained from the total Fe(II) concentration, provided that the $p\text{H}$ profile is known. The $p\text{H}$ is calculated by taking the chemical reactions listed in Table 1 into consideration. The concentrations of the major ions are assumed to be equal to the present concentrations, except for Ca^{2+} and SO_4^{2-} . The value $[\text{SO}_4^{2-}] = 10^{-3} \text{ mol kg}^{-1}$ is adopted, based on the discussion of Walker and Brindley¹¹. The Ca^{2+} concentration is chosen in such a way that the deep water is saturated with respect to calcite. $[\text{HCO}_3^-]$ and $[\text{CO}_3^{2-}]$ are calculated from the requirement that the surface water be in equilibrium with the atmospheric CO_2 pressure, which is a model parameter.

The Fe(II) concentration in the mixed layer is linked to its concentration in the deep water (at siderite saturation), by imposing that the upward flux of Fe(II) is conserved throughout the thermocline. The upward transport in this layer is due to advection and eddy diffusion. The diffusion coefficient K is assumed to be $4,000 \text{ m}^2 \text{ yr}^{-1}$, its present mean value¹². The advective velocity w is a model parameter; for the global ocean, its present value is $\sim 4 \text{ m yr}^{-1}$.

Other details of the model will be presented elsewhere (L.M.F. and J.-C. Gérard, in preparation), where chemical oxidation of Fe(II) by oxygen will be studied quantitatively and compared with photochemical oxidation.

The model has been run with atmospheric CO_2 pressures, P_{CO_2} , of 1, 10, 100 and 1,000 PAL (PAL, present atmospheric level). Table 2 lists the calculated values of ϕ_{phot} . The 'standard model' refers to the case $K = 4,000 \text{ m}^2 \text{ yr}^{-1}$ and $w = 4 \text{ m yr}^{-1}$; that is, it represents the ocean as a whole with conditions close to those pertaining at present. As the atmospheric CO_2 pressure is increased, two processes act in opposite directions. First, the $p\text{H}$ in the mixed layer becomes lower so that $[\text{FeOH}^+]$ and ϕ_{phot} tend to be reduced. Second, the carbonate ion becomes less abundant and, since the deep water is saturated with respect to siderite, the Fe^{2+} concentration increases in the deep water and also in the mixed layer as a result of transport, so that ϕ_{phot} tends to become higher. In the 'standard model', the second process is dominant and ϕ_{phot} increases from $0.308 \text{ mol Fe m}^{-2} \text{ yr}^{-1}$ at 1 PAL to $0.657 \text{ mol Fe m}^{-2} \text{ yr}^{-1}$ at 1,000 PAL. The effects of varying several parameters of the model have been investigated. Increasing the quantum efficiency to the value $\gamma = 0.05$ has no appreciable effect on ϕ_{phot} . Thus, the experimental uncertainty in the value of γ is not important. If the temperature of the deep water is increased from 5 to 20 °C, the rate of photo-oxidation is decreased as a result of lower Fe(II) concentrations caused by the lower solubility of siderite, but this decrease is not dramatic. The effect of dividing the absorption coefficient of FeOH^+ by 13 is marked at high P_{CO_2} . However,

Table 1 Reactions considered for the $p\text{H}$ calculation

	Reactions	log K	Ref.
K_1	$\text{Fe}^{2+} + \text{OH}^- \rightleftharpoons \text{FeOH}^+$	3.86	18
K_2	$\text{Fe}^{2+} + \text{Cl}^- \rightleftharpoons \text{FeCl}^+$	-0.45	18
K_3	$\text{Fe}^{2+} + \text{SO}_4^{2-} \rightleftharpoons \text{FeSO}_4^0$	0.61	18
K_4	$\text{H}^+ + \text{SO}_4^{2-} \rightleftharpoons \text{HSO}_4^-$	1.064	19
K_5	$\text{HCO}_3^- + \text{Mg}^{2+} \rightleftharpoons \text{MgHCO}_3^+$	0.017	19
K_6	$\text{CO}_3^{2-} + \text{Mg}^{2+} \rightleftharpoons \text{MgCO}_3^0$	1.512	19
K_7	$\text{HCO}_3^- + \text{Ca}^{2+} \rightleftharpoons \text{CaHCO}_3^+$	0.017	Assumed equal to K_5
K_8	$\text{CO}_3^{2-} + \text{Ca}^{2+} \rightleftharpoons \text{CaCO}_3^0$	1.512	Assumed equal to K_6
K_9	$\text{Mg}^{2+} + \text{OH}^- \rightleftharpoons \text{MgOH}^+$	1.58	19
K_{10}	$\text{Ca}^{2+} + \text{OH}^- \rightleftharpoons \text{CaOH}^+$	0.25	19
K_{11}	$\text{H}_2\text{O} \rightleftharpoons \text{H}^+ + \text{OH}^-$	*	20
K_{12}	$\text{H}_2\text{CO}_3 \rightleftharpoons \text{H}^+ + \text{HCO}_3^-$	*	21
K_{13}	$\text{HCO}_3^- \rightleftharpoons \text{H}^+ + \text{CO}_3^{2-}$	*	21
K_{14}	$\text{H}_3\text{BO}_3 \rightleftharpoons \text{H}^+ + \text{H}_2\text{BO}_3^-$	*	21

* K_{11} to K_{14} are temperature- and salinity-dependent expressions. A salinity of 35‰ has been assumed.

oxidized and gives rise to Fe(III) falling particles. In this model, it is assumed that no reduction of Fe(III) takes place in the thermocline, so that there is no production or sink of Fe(II) in this layer. Some Fe(III) may, nevertheless, be reduced to Fe(II) in the deep water. As a result, the rate of BIF deposition may be somewhat lower than the photo-oxidation rate ϕ_{phot} .

The model is composed of three coupled sub-models, the first of which describes the photochemical aspects of the photo-oxidation reaction, the second calculates the $p\text{H}$ profile and the third deals with ocean dynamics. These various calculations are repeated iteratively until convergence.

In the photochemical model, the photo-oxidation rate ϕ_{phot} is calculated as a function of the FeOH^+ concentration in the mixed layer. The solar photons are absorbed in the mixed layer by FeOH^+ and water. Photo-oxidation occurs for photons with a wavelength between 220 and 460 nm. The molar absorption coefficient $\epsilon_{\text{FeOH}^+}^+$ is from ref. 9. However, Braterman *et al.*¹⁰ report an absorption coefficient at 366 nm that is about 13 times less than that found by Ehrenfreund and Leibenguth⁹ at the same wavelength. Thus, the effect of dividing the absorption coefficient by this factor (over the whole spectrum) is assessed.

Table 2 The magnitude of the FeOH^+ photo-oxidation rate ϕ_{phot} at various atmospheric CO_2 pressures

ϕ_{phot} (mol Fe m^{-2} yr^{-1})	$P_{\text{CO}_2} =$ 1 PAL	$P_{\text{CO}_2} =$ 10 PAL	$P_{\text{CO}_2} =$ 100 PAL	$P_{\text{CO}_2} =$ 1,000 PAL
(1) Standard model	0.308	0.335	0.440	0.657
(2) $\gamma = 0.05$	0.309	0.337	0.444	0.676
(3) $T_{\text{DEEP}} = 20^\circ\text{C}$	0.203	0.220	0.280	0.406
(4) Modified absorption coefficient	0.301	0.312	0.359	0.361
(5) $T_{\text{DEEP}} = 20^\circ\text{C}$ + modified absorption coefficient	0.198	0.204	0.225	0.215
(6) $w = 0$	0.195	0.213	0.280	0.427
(7) $w = 0$ + modified absorption coefficient	0.192	0.203	0.245	0.278
(8) $w = 4,000$ m yr^{-1}	75.1	42.3	20.7	9.11

(1) Standard model: $K = 4,000$ $\text{m}^2 \text{yr}^{-1}$; $w = 4$ m yr^{-1} ; $\gamma = 0.03$; temperature of the deep water, 5°C . (2) Same as the standard model except that $\gamma = 0.05$. (3) Same as the standard model except that the temperature of the deep water is 20°C . (4) Same as the standard model except that the absorption coefficient $\varepsilon_{\text{FeOH}^+}$ of Ehrenfreund and Leibenguth⁹ has been divided by 13. (5) Same as model (4) except that the temperature of the deep water is 20°C . (6) Same as the standard model except that $w = 0$. (7) Same as model (4) except that $w = 0$. (8) Same as the standard model except that $w = 4,000$ m yr^{-1} .

even at 1,000 PAL, this variation of $\varepsilon_{\text{FeOH}^+}$ has no great effect on the photo-oxidation rate: ϕ_{phot} shows a decrease of only $<50\%$. Therefore, it seems that the mean photo-oxidation rate ϕ_{phot} for the global Archaean ocean could have been as high as 0.5 mol Fe $\text{m}^{-2} \text{yr}^{-1}$.

The effect of the intensity of the ocean circulation is also shown in Table 2. When w is set equal to zero, the model becomes entirely diffusive and the Fe(II) concentration increases linearly with depth. It is observed that the order of magnitude of ϕ_{phot} is not greatly affected by such a variation of w . On the other hand, when w is increased to $4,000$ m yr^{-1} , ϕ_{phot} can become several hundred times higher.

In the modern ocean, the vertical velocity of upwelling may be as large as several km yr^{-1} in areas of strong upwelling¹³. Thus, the model with $w = 4,000$ m yr^{-1} may be used to describe a restricted basin in such an area. At a CO_2 pressure of 100 PAL, the rate of photo-oxidation of FeOH^+ could have been as high as 20 mol Fe $\text{m}^{-2} \text{yr}^{-1}$, that is, 112 mg $\text{cm}^{-2} \text{yr}^{-1}$. This value may be compared with the iron deposition rate of 20 mg $\text{cm}^{-2} \text{yr}^{-1}$ in the BIFs of the Hamersley Basin in Australia¹⁴. Consequently, the possibility that extensive BIF deposits could form without oxygen through photo-oxidation of FeOH^+ must be considered.

If the area of the Archaean ocean was the same as today, a mean photo-oxidation rate of 0.5 mol Fe $\text{m}^{-2} \text{yr}^{-1}$ corresponds to a global flux of 1.8×10^{14} mol Fe yr^{-1} . Garrels and MacKenzie¹⁵ report a value of 12.2×10^{14} g yr^{-1} (that is, 0.218×10^{14} mol yr^{-1}) for the present-day input of iron to the ocean in the suspended sediments of rivers. On the Archaean Earth, however, the geochemical fluxes of ions into the ocean were probably governed by the interaction of sea water with the basaltic ocean crust¹⁶. For the present ocean, these fluxes are very poorly known, but according to Wolery and Sleep¹⁷, the total basaltic crust input of iron to the ocean would fall in the range $22\text{--}39 \times 10^{12}$ g $\text{Fe} \text{yr}^{-1}$; that is, $0.39\text{--}0.69 \times 10^{12}$ mol Fe yr^{-1} . This basaltic flux of iron was probably several times greater during the Archaean, as a result of a higher rate of volcanic activity. However, this flux was probably not much greater than the present flux of iron suspended in rivers. Thus, a conservative estimate is that the Archaean geochemical input of iron to the ocean was close to 0.2×10^{14} mol yr^{-1} . If this assumption is correct, the calculated value of the global photo-oxidation rate of FeOH^+ , ϕ_{phot} , seems to be an order of magnitude too high.

This discrepancy may be explained by one or more of the following hypotheses: (1) The deep water was not saturated with respect to siderite. (2) The absorption of solar light by sea salts in the range $220\text{--}460$ nm can limit the photo-oxidation rate more efficiently than the absorption by water itself. (3) The Fe(III) produced in the mixed layer reacted with reduced compounds dissolved in the ocean, thus decreasing the rate of BIF deposition. (4) The diffusion coefficient K and the advective velocity w were both lower than today.

It would be interesting to refine the ocean model to study possibilities (2) and (3). However, such a task would not be easy in view of our poor knowledge of the Archaean ocean chemistry.

In conclusion, the present model shows that FeOH^+ photo-oxidation was able to produce BIF deposits as extensive as have been found in the Hamersley Basin. From a global point of view, a total photo-oxidation rate of 1.8×10^{14} mol Fe yr^{-1} has been calculated. Since FeOH^+ photo-oxidation releases hydrogen, such a flux may have had important consequences for the oxidation state of the Precambrian atmosphere. An important question concerns the rate at which Fe(III) would react with the hydrogen released or with any other reduced compound dissolved in the ocean. This point is difficult to analyse quantitatively as neither the kinetics of the reaction nor the composition of Archaean sea water is well known. The model results might also depend slightly on the presence of small amounts of O_2 . This problem will be examined elsewhere.

This work was supported by a fellowship from the Belgian Institute for the encouragement of Industrial and Agricultural Scientific Research.

Received 19 September 1985; accepted 20 January 1986.

- Gole, M. J. & Klein, C. *J. Geol.* **89**, 169–183 (1981).
- James, H. L. *U.S. geol. Surv. Prof. Pap.* **440-W** (1966).
- Holland, H. D. *Econ. Geol.* **68**, 1169–1172 (1973).
- Drever, J. I. *Geol. Soc. Am. Bull.* **85**, 1099–1106 (1974).
- Walker, J. C. G. *Pure appl. Geophys.* **117**, 498–512 (1978).
- Cairns-Smith, A. G. *Nature* **276**, 807–808 (1978).
- Braterman, P. S., Cairns-Smith, A. G. & Sloper, R. W. *Nature* **303**, 163–164 (1983).
- Langmuir, D. *U.S. Geol. Surv. Prof. Pap.* **650-B**, 180–184 (1969).
- Ehrenfreund, M. & Leibenguth, J.-L. *Bull. Soc. chim. Fr.* 2494–2505 (1970).
- Braterman, P. S., Cairns-Smith, A. G., Sloper, R. W., Truscott, T. G. & Craw, M. *Dalton Trans.*, 1441–1445 (1984).
- Walker, J. C. G. & Brimblecombe, P. *Precamb. Res.* **28**, 205–222 (1985).
- Oeschger, H., Siegenthaler, U., Schotterer, U. & Gugelmann, A. *Tellus* **27**, 168–192 (1975).
- Holland, H. D. *The Chemical Evolution of the Atmosphere and Oceans*, 396 (Princeton University Press, 1984).
- Trendall, A. F. *J. geol. Soc. Aust.* **19**, 287–311 (1972).
- Garrels, R. M. & Mackenzie, F. T. *Evolution of Sedimentary Rocks*, 112 (Norton, New York, 1971).
- Veizer, J., Compston, W., Hoefs, J. & Nielsen, H. *Naturwissenschaften* **69**, 173–180 (1984).
- Wolery, T. J. & Sleep, N. H. *J. Geol.* **84**, 249–275 (1976).
- Davison, W. *Geochim. cosmochim. Acta* **43**, 1693–1696 (1979).
- Dyrssen, D. & Wedborg, M. in *The Sea Vol. 5* (ed. Goldberg, E. D.) 181–195 (Wiley, New York, 1974).
- Millero, F. J. *Geochim. cosmochim. Acta* **43**, 1651–1661 (1979).
- Gieskes, J. M. in *The Sea Vol. 5* (ed. Goldberg, E. D.) 123–151 (Wiley, New York, 1974).

High levels of natural radionuclides in a deep-sea infaunal xenophyphore

David D. Swinbanks & Yoshihisa Shirayama

Ocean Research Institute, University of Tokyo, 1-15-1 Minamidai, Nakano-ku, Tokyo 164, Japan

While debate continues about the disposal of man-made radioactive wastes in the deep sea, one deep-sea rhizopod is storing naturally occurring radionuclides to high concentrations in its home centimetres below the floor of the Izu–Ogasawara Trench. The hadal infaunal xenophyphore *Oculitamina profunda* fills its test, a network of sediment tubes, with spherical micro-packages of waste (stercomes) containing $450\text{--}520$ d.p.m. per g dry wt of the natural radionuclide ^{210}Pb . The stercomes are stored in flimsy membrane tubes immediately next to a thread-like protoplasmic body (granellare), which has equally high ^{210}Pb levels. A model

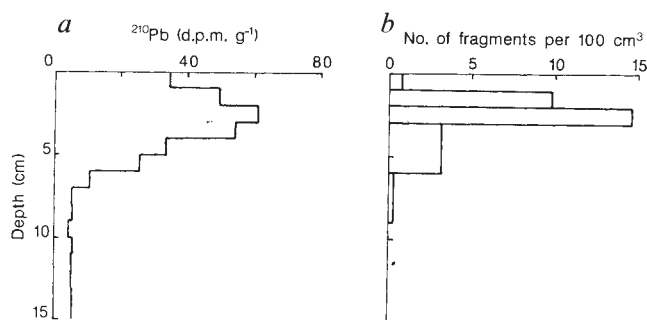


Fig. 1 The vertical distribution of ^{210}Pb (ref. 5) (a) and fragments of *O. profunda*⁸ (b) in the box core of sediment from the Izu-Ogasawara Trench (sample depth 8,260 m). The former distribution is based on one subcore of 100 cm² cross-section while the latter is the average distribution of four such subcores (not including that for ^{210}Pb). Only fragments of *O. profunda* which contained granellare stained red by Rose Bengal were counted. Each fragment contains on average $2.4 \pm 2.3 \mu\text{g}$ dry wt of granellare ($n = 42$).

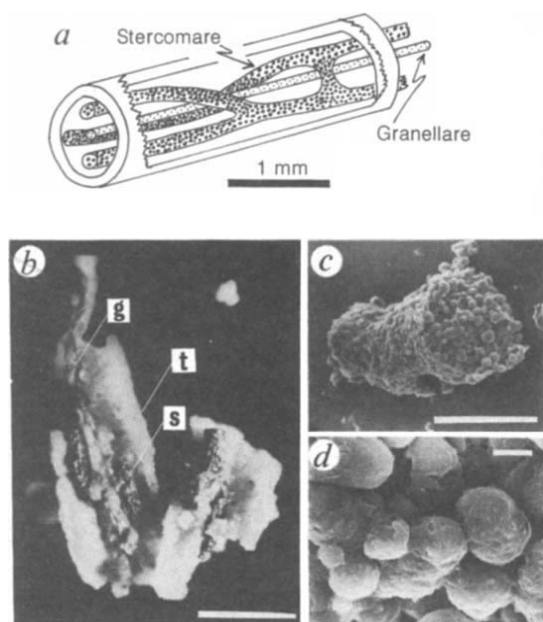


Fig. 2 a, Internal organization of *O. profunda*. b, A fragment of *O. profunda* split open to reveal the granellare (g), stercomare (s) and tubular test (t). Scale bar, 500 μm . c, Scanning electron micrograph of a stercomare showing aggregations of stercomeres (pellets) enclosed within a tubular membrane. Scale bar, 170 μm . d, Scanning electron micrograph of stercomeres. Scale bar, 8 μm .

of stercome excretion suggests that *O. profunda* grows quickly and can produce transient subsurface peaks in the vertical distribution of ^{210}Pb in the sediment while causing little, if any, sediment mixing, a finding which could have profound consequences for the ^{210}Pb -modelling of deep-sea bioturbation (mixing of sediments by organisms). From the barium content of the granellare (21,000 parts per million, p.p.m.) and assuming that the $^{226}\text{Ra}/\text{Ba}$ ratio in the surrounding sediment and sea water is maintained in the organism, it is inferred that the protoplasmic ^{210}Pb is supported by 320–350 d.p.m. per g dry wt of ^{226}Ra concentrated in intracellular barite crystals. Consequently, *O. profunda* and xenophyophores in general are probably subject to unusually high levels of natural radiation (several Sv yr^{-1}), among the highest reported for any living organism.

^{210}Pb is derived from the decay of ^{226}Ra in sea water and the Earth's crust¹. With a half-life of 22 yr, ^{210}Pb decays to the short-lived β -emitter ^{210}Bi and thence to the α -emitter ^{210}Po which has a half-life of 138 days. ^{210}Po is often concentrated by marine organisms and unusually high levels of ^{210}Po have

recently been reported in certain mid-water oceanic shrimp and fish, but ^{210}Pb activities in organisms are typically much lower (a few d.p.m. per g dry wt at most)². However, very little is known about natural radionuclide levels in deep-sea benthic organisms, a situation that must be rectified if the hazards of radioactive waste disposal are to be properly assessed².

The vertical distribution of ^{210}Pb in surficial deep-sea sediments has been analysed extensively in attempts to estimate rates of bioturbation^{3–7}. The mixing model on which such estimates are based, however, assumes that ^{210}Pb enters the sediment at the sediment/water interface attached to sinking particulate matter and remains attached to sedimentary particles throughout the mixing process. The possibility that a significant fraction of ^{210}Pb enters the sediment through the tissues and excretion products of benthic organisms living below the sediment/water interface has not been considered. We decided to examine ^{210}Pb levels in the infaunal xenophyophore *O. profunda* because an unusual subsurface peak in the vertical distribution of ^{210}Pb in the sediment, which could not be explained by the standard mixing model⁵, coincided with a peak in the vertical distribution of the xenophyophore at 2–3 cm depth⁸, and we thought the xenophyophore might be responsible for the ^{210}Pb peak (Fig. 1). Such 'anomalous' ^{210}Pb distributions can be quite common in the deep sea⁶, and it is becoming increasingly clear that the standard mixing model, which predicts an exponential decrease in ^{210}Pb activity with depth in the sediment, is 'idealized'⁷.

Xenophyophores are giant deep-sea rhizopod protozoans that live on and in the sediment^{8,9} and are closely related to benthic foraminiferans. Deep-sea photography has revealed that epifaunal xenophyophores can be very abundant ($\sim 1\text{--}10 \text{ m}^{-2}$) in certain environments¹⁰, and the infaunal xenophyophore of the present study, collected from a box core from the Izu-Ogasawara Trench, was the most abundant macrofaunal organism in the core⁸.

The test of *O. profunda* consists of a horizontal network of sediment tubes (500–1,000 μm diameter) that enclose a central granellare string (40–130 μm diameter) surrounded by one to three peripheral stercomare strings (125–160 μm diameter) (Fig. 2a–c)⁸. Both the granellare and stercomare have a thin ($<0.5 \mu\text{m}$) organic membrane wall that in the former encloses the multinucleate plasma body and in the latter, stercomes. Stercomes are small black spherical pellets $\sim 10 \mu\text{m}$ in diameter (Fig. 2d) which are thought to be waste and excretion products, although very little is known about their composition⁹. An unusual feature of xenophyophores is that they contain high concentrations of barium ($\sim 10^4$ p.p.m.) in the form of intracellular barite crystals called granellae, which are believed to be secreted by the organism^{9,11}. In *O. profunda* the granellae are less than 0.5 μm in size (O. S. Tendal, personal communication).

A total of about 160 fragments of *O. profunda* were separated from two subcores of sediment (each 100 cm² in horizontal cross-section) which had been stored in a freezer (-20°C) since collection in March 1980. The top 10 cm of each subcore was sliced horizontally at 1-cm intervals and the sediment washed through a 0.5-mm sieve that retained fragments of the organism. After air-drying on filter paper, the fragments were divided into test, granellare and stercomare, each subcore yielding ~ 40 mg dry weight of test, 5 mg of stercomare and 0.2 mg of granellare. In addition, comparable amounts of test, stercomare and granellare were obtained from alcohol-preserved specimens of *O. profunda*. Pooled samples of each part of the organism and the sediment at each 1-cm interval were analysed separately for ^{210}Pb by the ^{210}Po technique¹² (Table 1). ^{210}Pb activities were corrected to those at the time of sampling (4–5 yr earlier), assuming that the ^{210}Pb (half-life 22 yr) was unsupported. Aliquots (10 cm³) of the acid-dissolved samples (50 cm³) used for ^{210}Pb analysis were analysed for barium by inductively coupled plasma emission spectrophotometry¹³.

The ^{210}Pb content of the sediment peaked at 1–2 cm depth, coincident with a peak in the vertical distribution of *O. profunda*,

Table 1 Measured ^{210}Po activities and barium contents of various parts of *O. profunda* and the surrounding sediment, and their estimated ^{210}Pb and ^{226}Ra activities

Sample	^{210}Po (d.p.m. g $^{-1}$)*	^{210}Pb (d.p.m. g $^{-1}$)	Ba (p.p.m.)	^{226}Ra (d.p.m. g $^{-1}$)
Granellare (0–6 cm)	450 ± 100†	510 ± 110	21,000	320–350
Stercomare (0–6 cm)	380 ± 60‡	450 ± 70	870	13–14
1–2 cm	440 ± 50§	490 ± 50		
2–3 cm	420 ± 40†	480 ± 50		
Test (0–6 cm)	440 ± 50‡	520 ± 60	356	5–6
2–3 cm	430 ± 40‡	510 ± 50		
Sediment (0–6 cm)	42 ± 5	47 ± 6		
	52 ± 3	62 ± 4		
	9–57	10–65	329–359	5.2¶

Samples were digested in acid and ^{210}Po was spontaneously deposited on one or two silver disks (1 cm diameter) the α -activities of which were then counted for 1–5 days. A ^{208}Po tracer was added to the sample for the determination of chemical yield, which ranged between 10 and 85%. Because the ^{210}Po analysis was performed 4–5 yr after sample collection, it can be assumed that ^{210}Po , which has a half-life of 138 days, was in secular equilibrium with its grandparent ^{210}Pb at the time of analysis (that is, the activities of the two radionuclides were the same) and this was confirmed by replating some of the samples after 9–12 months. Errors are based on 1σ counting statistics and include, where significant, errors due to blank correction. Counts for chemical blanks were at most one or two counts per day above background, which at counting efficiencies of 24–33% ranged from 0.3 to 7 counts per day depending on the detector used (eight detectors in all). ^{210}Po counts for the granellare samples, although low due to low sample weight, were more than five times those for the blanks (including background), yielding ^{210}Po estimates with a 15–20% error as opposed to ~10% error for the other parts of the organism.

* 1 d.p.m. g $^{-1}$ = 16.7 Bq kg $^{-1}$.

† Measured 26 March 1984 (frozen samples, collected 3 March 1980).

‡ Measured 27 September 1985 (alcohol-preserved samples, collected 3 March 1980).

§ Measured 15 December 1983 (frozen samples, collected 3 March 1980).

|| Estimated from Ba content assuming a $^{226}\text{Ra}/\text{Ba}$ ratio of 4.2–4.6 nmol per mol.

¶ Estimated from constant ^{210}Pb activity below 7 cm depth (Fig. 1).

but the ^{210}Pb peak was much weaker and 1 cm shallower than in Fig. 1 (Fig. 3a, b). The ^{210}Pb activity of the xenophyophore's test, which is composed of particles selected from the surrounding sediment, was comparable to, or slightly higher than, the activity of sediment at the same depth (Fig. 3a), whereas the activity of the granellare and stercomare was an order of magnitude higher, at 450–520 d.p.m. per g dry wt (Table 1). In the 1–2- and 2–3-cm layers, where *O. profunda* was most abundant, stercomare showed no significant difference in ^{210}Pb content between layers, with values of 520 and 510 d.p.m. per g dry wt, respectively. Comparably high levels of unsupported ^{210}Po have recently been reported in the hepatopancreas of mid-water penaeid shrimp, but the highest levels of ^{210}Pb previously reported in any organism, which were for benthic amphipods from the Central Pacific, are two orders of magnitude lower than our findings². Preliminary ^{210}Pb analysis of two epifaunal species of xenophyophore (an unidentified species from a depth of 5,270 m on the Pacific margin of the Japan Trench and *Psammina nummulina* collected from 3,570 m between the Cocos Ridge and East Guatemala Basin) has yielded comparably high levels of ^{210}Po in their granellare and stercomare, but ^{210}Pb support has yet to be confirmed (our unpublished data).

There are two possible origins for the ^{210}Pb in *O. profunda*. One possibility is that the xenophyophore accumulates the radionuclide at the sediment/water interface by feeding on settling particulate matter containing several hundred d.p.m. per g dry wt of ^{210}Pb (refs 14, 15; for a possible feeding mechanism

in infaunal xenophyophores see ref. 16). Another quite different possible source is ^{226}Ra within the intracellular granellae. It is well known that the chemistry of barium is very similar to that of radium, its chemical congener in the Periodic Table, and the $^{226}\text{Ra}/\text{Ba}$ ratio in oceanic sea water maintains a nearly constant value of 4.6 nmol of ^{226}Ra per mol Ba (ref. 17), very similar to the value estimated for the box-core sediment of the present study (we estimate the ratio to be 4.2 nmol ^{226}Ra per mol Ba by assuming that the almost constant ^{210}Pb activity of about 5.2 d.p.m. per g dry wt below 7-cm depth (Fig. 1) equals the ^{226}Ra activity and by using an average of 335 p.p.m. for the barium content of the sediment (Table 1)). Assuming that the ratio in the surrounding sediment and sea water is maintained in *O. profunda*, as in diatoms¹⁸, the granellare, which contain about 21,000 p.p.m. of barium (Table 1), are estimated to contain 320–350 d.p.m. per g dry wt of ^{226}Ra (using values of 4.2 and 4.6 nmol ^{226}Ra per mol Ba, respectively). If this estimate is of the right order of magnitude, then estimated ^{210}Pb contents for granellare in Table 1 will require a downward adjustment of ~10% (assuming no radon loss) because in the calculation of the values in Table 1 it was assumed that ^{210}Pb was unsupported.

The percentage of ^{210}Pb in the granellare which is derived from ^{226}Ra in the granellae will depend on the age of the granellae. If, as we suggest below, xenophyophores are short-lived organisms, and if granellae are secreted intracellularly, then the contribution of ^{210}Pb from granellae will be small, because it requires several half-lives (>100 yr) for ^{210}Pb to achieve secular equilibrium with ^{226}Ra . In the case of the stercomare and test, the barium content is much lower, 870 and 356 p.p.m., respectively, corresponding to estimated ^{226}Ra contents of 13–14 and 5–6 d.p.m. per g dry wt, respectively (Table 1), and thus their ^{210}Pb is largely unsupported by ^{226}Ra .

Our ^{210}Po data and the ^{226}Ra levels estimated from the barium content suggest that the plasma body of *O. profunda* is subject to an unusually high dose of natural radiation. Considering only α -radiation and assuming that it has an effective quality factor of 20 (ref. 19), the dose from ^{210}Pb -supported ^{210}Po alone is about 0.8 Sv yr $^{-1}$ (using a measured wet/dry ratio of 5.0 ± 0.5 and a ^{210}Po (^{210}Pb) content of 450 d.p.m. per g dry wt), whereas if the granellare contain 320 d.p.m. per g dry wt of ^{226}Ra (the lower estimate given above) the radiation dose increases to about 3.0 Sv yr $^{-1}$ (assuming no radon loss and assuming the short-lived α -emitting descendants of radon, namely ^{218}Po and ^{214}Po , contribute to the dose). The latter dose is comparable to that in the hepatopancreas of mid-water penaeid shrimp (3.9 Sv yr $^{-1}$), which is the highest natural radiation dose so far reported for any living organism². The dose from ^{226}Ra for *O. profunda*, however, should be regarded only as an order of magnitude estimate in the absence of measured ^{226}Ra levels.

In most xenophyophores, the granellae are much larger (2–5 μm) than in *O. profunda* (<0.5 μm) and in some species granellae can be so abundant as to obscure the nuclei of the plasma body⁹. If the $^{226}\text{Ra}/\text{Ba}$ ratio of oceanic sea water is maintained in these crystals, they will contain about 10^4 d.p.m. g dry wt of ^{226}Ra and the adjacent nuclei of the plasma body will be subject to radiation doses even higher than those estimated above for *O. profunda*.

Although marine organisms are probably less sensitive to radiation than is man²⁰, the possibility that these high radiation doses have affected xenophyophores over the long term should be considered. For example, has exposure to high levels of radiation resulted in a high incidence of radiation-induced mutation and/or development of radiation resistance? Studies of the molecular biology of these organisms might answer some of these questions. Turning to the geological record, it has been suggested that the trace fossil *Paleodictyon*, which occurs as far back as the Ordovician, is a fossilized form of infaunal xenophyophore¹⁶. If this is so, study of *Paleodictyon* could throw light on the evolution of xenophyophores over the past 500 Myr and might provide further insight into the above matters.

Fig. 3 *a*, Activity of excess ^{210}Pb in the sediment ($\square$, $\blacksquare$) and test ($\circ$) of *O. profunda* as a function of sediment depth. Subcores B ($\blacksquare$) and E ($\square$) were separated laterally by 10 cm. All the test samples come from subcore E and covered a 1-cm vertical interval except for that from 3 to 6 cm, which came from a separate subcore not analysed for ^{210}Pb . Horizontal bars indicate errors in test values. Excess ^{210}Pb was calculated by

assuming a constant ^{226}Ra content of 5.2 d.p.m. per g in the sediment and test (estimated from constant ^{210}Pb activity below 7 cm depth in Fig. 1 and barium content of test in Table 1). *b*, Average vertical distribution of *O. profunda* stercomare (in weight per cent of total stercomare yield) in the two subcores in *a*. *c*, A model of the effects of stercomare excretion on the vertical distribution of ^{210}Pb . The sediment is divided into 1-cm-thick horizontal layers with a horizontal cross-sectional area of 100 cm² (as in our subcores) and the activity of excess ^{210}Pb is initially assumed to decrease exponentially with depth, as is often the case in deep-sea sediments. The inventory of excess ^{210}Pb is maintained at 60 d.p.m. cm⁻² (as in our subcores) by a constant flux of ^{210}Pb to the sediment/water interface. For simplicity, porosity is assumed to be constant with depth (the average value of 86.5% for 0–7 cm is used) and sedimentation is considered to be negligible during the short time intervals considered. The concentration of excess ^{210}Pb in stercomare is taken to be constant at 500 d.p.m. g⁻¹ and the stercomare ^{210}Pb is assumed to be derived from feeding at the sediment/water interface (all the ^{210}Pb in stercomare comes from the 0–1-cm layer). In the model, a given amount of stercomare (5, 10 or 20 mg per 100 cm² for cases 1, 2 and 3, respectively) is instantaneously formed each month while an equivalent amount decays and releases its ^{210}Pb to the immediately surrounding sediment. After stercomare decay, it is assumed that the stercomare-derived ^{210}Pb is not immediately redispersed by bioturbation, an assumption which probably holds for periods of time comparable to the half-life of ^{210}Pb (22 yr) (ref. 6). If the rate of stercomare formation is r , then the total inventory of stercomare-derived ^{210}Pb after time t will be given by: $I_t = \int_0^t ar e^{-\lambda t} dt = ar(1/\lambda - 1/\lambda e^{-\lambda t})$, where λ is the decay constant of ^{210}Pb and a is the activity of ^{210}Pb in stercomare. The inventory is subtracted from the 0–1-cm layer and portions are allotted to the layers of the model according to the depth distribution of stercomare on the left. The effects of the three rates of stercomare formation in cases 1, 2 and 3 are considered. The effects of these rates after various intervals of time are shown by the dashed and solid curves ($\bullet$, Exponential curve; $\blacktriangle$, case 1, 250 yr; case 2, 22 yr; case 3, 9 yr; $\triangle$, case 2, 250 yr; case 3, 22 yr). For example, the dashed curve can be produced by a rate of 10 mg per 100 cm² per month (case 2) acting for 22 yr or by a rate of 20 mg per 100 cm² per month (case 3) acting for 9 yr, while the solid curve can be produced by the latter rate acting for 22 yr. Note the similarity of the dashed curve to the curves in *a* and the resemblance of the solid curve to the ^{210}Pb profile in Fig. 1.

Of more immediate interest, however, is the possible effect of *O. profunda* on ^{210}Pb distribution in the sediment through the excretion of stercomes rich in unsupported ^{210}Pb . While it is generally believed that deep-sea organisms grow slowly and live long, there is circumstantial evidence that xenophyophores may grow rapidly. Epifaunal xenophyophores have been observed on freshly formed biogenic mounds (L. A. Levin, personal communication) which elsewhere in the deep sea have been seen to form within a few months (C. R. Smith, personal communication). As xenophyophores are almost certainly immobile, this suggests they may grow on a timescale of months rather than years. If so, a dense patch of *O. profunda* can have a very significant localized effect on the vertical distribution of ^{210}Pb , as the model in Fig. 3c shows. In this model, we consider only stercomare excretion—the contribution of ^{210}Pb from granelare, the weight of which is only 3–4% that of stercomare, is neglected—and it is assumed that all unsupported ^{210}Pb in stercomare is derived from feeding on particulate matter settling at the sediment/water interface. The particulate matter flux will thus mark an upper limit for the rate of stercomare production. We estimate the former to be ~50–60 mg per 100 cm² per month based on a ^{210}Pb flux of 1.9–2.4 d.p.m. per cm² per yr (ref. 5 and the present study) and an average ^{210}Pb content for particulate matter of 330 d.p.m. per g (refs 14, 15). At the maximum observed density of 420 *O. profunda* fragments per 100 cm², the density of stercomare is estimated to be ~30 mg per 100 cm², and if this is replaced at a rate between once a month and once a year, local stercomare production will lie between 2.5 and 30 mg per 100 cm² per month, well within the above estimated flux of particulate matter.

When stercomare production exceeds ~5 mg per 100 cm² per month, the vertical distribution of ^{210}Pb begins to deviate significantly from the exponential curve expected in the standard diffusion mixing model³ (compare with Fig. 3c). At a rate of 10 mg per 100 cm² per month, a small subsurface peak comparable to that observed in our subcores is formed within 22 yr at the depth of the peak in stercomare distribution (Fig. 3a–c); at a rate of 20 mg per 100 cm² per month the same peak is formed

in only 9 yr, after 22 yr it becomes pronounced (compare with Fig. 1) and after 40 yr the 0–1-cm layer is completely lacking in ^{210}Pb . Thus, a combination of patchy *O. profunda* distribution and rapid stercomare formation could account for both the occurrence of a subsurface ^{210}Pb peak and the marked subcore variation in the vertical distribution of ^{210}Pb . Looked at another way, our model and the observed distributions of ^{210}Pb and *O. profunda* suggest that this xenophyophore grows and excretes rapidly.

The model in Fig. 3c represents a major departure from the usual interpretation of ^{210}Pb distribution in deep-sea sediments, as it involves highly selective redistribution of ^{210}Pb -rich material in a non-steady-state fashion with little or no sediment mixing (our observations of stercomes suggest they contain few if any refractory sedimentary minerals). Apart from xenophyophores, many 'primitive' monothalamous deep-sea foraminiferans (for example, allogromiids, *Rhizammina*, komokiaceans) accumulate stercomes, and other benthic rhizopod protozoans producing masses of stercomes are extremely abundant in oligotrophic regions of the deep sea²¹. Our model therefore probably applies to wide areas of the ocean floor, and ^{210}Pb distributions in deep-sea sediments may require substantial reinterpretation.

We thank N. Nozaki for providing advice and facilities for ^{210}Po analysis, H. S. Yang and M. Yamada for assistance in radionuclide analysis, M. Sugiyama for carrying out the barium analysis, and R. D. Cherry, J. N. Smith, A. J. Gooday and J. Thomson for their review comments. D.D.S. was supported by a postdoctoral fellowship from the Japan Society for the Promotion of Science and the Royal Society of London.

Received 30 December 1985; accepted 28 January 1986.

- Burton, W. M. & Stewart, N. G. *Nature* **186**, 584–589 (1960).
- Cherry, R. D. & Heyraud, M. *Science* **218**, 54–56 (1982).
- Nozaki, Y., Cochran, J. K., Turekian, K. K. & Keller, G. *Earth planet. Sci. Lett.* **34**, 167–173 (1977).
- DeMaster, D. J. & Cochran, J. K. *Earth planet. Sci. Lett.* **61**, 257–271 (1982).
- Yamada, M., Kitaoka, H. & Tsunogai, S. *Deep Sea Res.* **30**, 1147–1156 (1983).
- Smith, J. N. & Schafer, C. T. *J. mar. Res.* **42**, 1117–1145 (1984).
- Cochran, J. K. *Geochim. cosmochim. Acta* **49**, 1195–1210 (1985).
- Tendal, O. S., Swinbanks, D. D. & Shirayama, Y. *Oceanol. Acta* **5**, 325–329 (1982).
- Tendal, O. S. *Galathea Rep.* **12**, 7–99 (1972).

10. Tendal, O. S. & Gooday, A. J. *Oceanol. Acta* **4**, 415-422 (1981).
11. Gooday, A. J. & Nott, J. A. *J. mar. biol. Ass. U.K.* **62**, 595-605 (1982).
12. Black, S. C. *Hum. Phys.* **7**, 87-91 (1961).
13. Sugiyama, M., Matsui, M. & Nakayama, E. *J. oceanogr. Soc. Jap.* **40**, 295-302 (1984).
14. Spencer, D. W. *et al. J. mar. Res.* **36**, 493-523 (1978).
15. Bacon, M. P., Huh, C., Fleer, A. P. & Deuser, W. G. *Deep Sea Res.* **32**, 273-286 (1985).
16. Swinbanks, D. D. *Science* **218**, 47-49 (1982).
17. Chan, L. H. *et al. Earth planet. Sci. Lett.* **32**, 258-267 (1976).
18. Szabo, B. J. *Geochim. cosmochim. Acta* **31**, 1321-1331 (1967).
19. International Commission on Radiological Protection, Publication **26**, 53 (Pergamon, Oxford, 1977).
20. Blaylock, B. G. & Trabalka, J. B. *Adv. Radiat. Biol.* **7**, 103-152 (1978).
21. Riemann, F. *Oceanol. Acta* **6**, 303-311 (1983).

Perception of apparent motion by commissurotomy patients

V. S. Ramachandran

Department of Psychology, C-009, University of California at San Diego, La Jolla, California 92093, USA

A. Cronin-Golomb & J. J. Myers

Division of Biology, California Institute of Technology, Pasadena, California 91125, USA

When two spatially separated light spots are flashed in rapid succession, the spot will appear to move between the two locations—an illusion called apparent motion^{1,2}. We have presented this display to callosum-sectioned human patients and found that they could correctly report the temporal order of a simple apparent motion sequence presented across the vertical meridian. Hence, the forebrain commissures are not required for this function.

We investigated motion perception by commissurotomy patients whose corpus callosum, anterior commissure, hippocampal commissure and massa intermedia (when encountered) had been sectioned surgically for the treatment of intractable epilepsy^{3,4}. An intact corpus callosum is presumed to be involved in various types of visual integration across the vertical meridian^{5,6}. For example, we tested three callosum-sectioned patients (L.B., A.A. and N.G.) and found that they experienced difficulty in matching the orientations of two lines flashed simultaneously on either side of the vertical meridian for 150 ms (ref. 7). When asked to indicate whether the orientations were the same or different, they performed very poorly; even orthogonal lines were sometimes reported to be parallel. Further, callosum-sectioned patients cannot perceive 'midline-stereopsis'⁸ because they lack the commissural fibres⁹⁻¹² required to match disparate binocular stimuli across the vertical meridian. We therefore wondered whether these individuals could perceive apparent motion of a single light spot jumping across the midline from one hemi-field to the other. The spot subtended 20 arc min and its two successive locations were 9° apart (that is, 4.5° from the midline), so that we were stimulating the long-range rather than short-range motion¹³ system. The spot was generated on a Panasonic monitor using an Apple 2c microcomputer. Stimulus duration and inter-stimulus interval (ISI) were held constant at 130 ms and 30 ms, respectively. Trials were generated randomly by the computer so that the spot jumped from either left to right (L-R) or right to left (R-L), or the two spots appeared simultaneously. One experimenter sat behind the cathode ray tube (CRT) screen and carefully watched the subject's eyes, and any trial on which the eyes moved was discarded. Each trial began a short interval after a warning tone. The interval between the tone and the appearance of the first spot was varied randomly. The subject's task was to fixate a central X and to indicate either the presence of simultaneity or the direction of perceived apparent movement by pointing to a card which pictorially depicted the three possibilities (Fig. 1). To ensure that the instructions were clearly understood, we began with 10 'training' trials presented entirely within the appropriate hemi-field alone

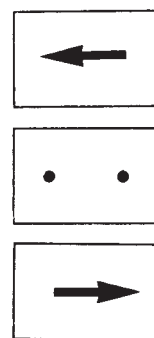


Fig. 1 The illustration shown to the subjects to enable them to point to the perceived direction of movement. The card was always presented in the hemi-field corresponding to the hand that was being used.

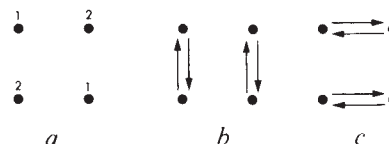


Fig. 2 Fluctuation in perception similar to that seen in Necker cubes can also be observed in an ambiguous apparent motion display (a). The direction of perceived oscillation in this bi-stable display can be influenced by the will only at slow speeds of alternation. Dots in diagonally opposite corners are flashed simultaneously and then switched off and replaced by dots appearing on the remaining two corners. (The numerals indicate order of presentation.) The frames are alternated in a continuous cycle. The two possible percepts (b and c) are also depicted.

(for example, left hemi-field for the left hand). When the actual experimental session began, the subject used one hand alone to indicate the perceived direction for each of 30 trials presented across the vertical meridian in random order. (The card with illustrations was always presented in the hemi-field corresponding to the hand that was being used.) After this, the subject switched hands and was confronted with a new random sequence of 30 trials. Three subjects ($n = 3$) were used in these experiments and in two of them (L.B. and N.G.) we had NMR confirmation of commissurotomy¹⁴ (J. Bogen, personal communication).

The results were clear-cut. L.B. could clearly discriminate L-R from R-L and either of these from simultaneous stimuli. (He was 100% accurate on all 60 trials, 30 with each hand.) N.G. had no difficulty in discriminating L-R from R-L, but she displayed considerable uncertainty with the simultaneous condition. (She was correct on 37 out of the 40 L-R/R-L trials and incorrect on all 20 trials on which the dots appeared simultaneously.) The third subject (A.A.) was accurate on 43 out of 60 trials and had no special difficulty with simultaneous stimuli (in fact, he got 18 correct out of 20). For all three subjects $P < 0.01$ (two-tailed binomial test). The result with L.B. is especially interesting because previous studies⁸ have shown him to lack midline stereopsis.

We were especially concerned about eye movements in this experiment, as the subject could have resorted to the strategy of using eccentric fixation to confine both spots to one hemi-field. To minimize this possibility, we instructed the subject to carefully fixate the central X, and had one experimenter constantly watching the eyes from across the table. As a further precaution, we ran another series of trials with the spots separated by 14° of visual angle instead of 9°. Again, L.B. was accurate on 100% of 30 trials, N.G. was correct on all 20 trials requiring R-L/L-R discrimination and incorrect on all 10 trials on which the spots were simultaneous, and A.A. was correct on 24 trials out of 30 ($P < 0.01$; two-tailed binomial test). In this experiment, the subjects would have had to deviate their fixation by at least 7°

to bring both spots within one hemi-field and this seems highly unlikely.

Thus, the forebrain commissures are not required for correctly discriminating the temporal order of stimuli appearing across the midline in rapid succession. However, were the patients actually seeing movement or were they simply making discriminations based on temporal order *per se*? We questioned them on this point, taking care not to prompt them, and all three replied unhesitatingly that they had seen the dot 'move' or 'jump' rather than merely change location. To confirm this further, we increased the ISI gradually until the subjects (L.B. and A.A.; N.G. was not tested) reported a transition from seeing movement to seeing simple temporal succession. (The duration of each spot was 50 ms and the separation between them 8°.) For A.A. the transition occurred at 456 ms (mean of 8 readings, 4 ascending and 4 descending; s.d. = 46) and for L.B. it occurred at 329 ms (s.d. = 28) for vertical apparent motion along the vertical meridian. For horizontal motion across the vertical meridian, these values were 372 ms (s.d. = 71.5) and 346 ms (s.d. = 49), respectively. Thus, the transition from motion to temporal succession seems to occur at about the same ISI for both vertical and horizontal sequence, suggesting that the patients were indeed seeing movement across the midline when the ISI was appropriate. Interestingly, for the horizontal sequence, when the ISI was too long, they often reported seeing only a single spot flashing on and off and their ability to discriminate temporal order actually deteriorated even though they had more time available. It was almost as though one could switch from a single conscious individual to two individuals by simply increasing the ISI! (This provides another suggestive argument against the involvement of eye movements.) Further experiments are in progress to determine whether Korte's laws¹⁴ of apparent motion would also apply to the 'second' visual system.

We conclude that apparent motion can be seen readily across the midline in spite of total transection of the forebrain commissures; this implies that the effect might be mediated largely by the so-called second visual system¹⁵⁻¹⁹, which bypasses the lateral geniculate nucleus and relays through the superior colliculus and pulvinar. This phylogenetically older system may be adequate for mediating motion perception but not for stereopsis²⁰⁻²² and 'form' perception. But what about more complex effects in motion perception, such as the perception of the ambiguous display depicted in Fig. 2? Does the bi-stability of ambiguous apparent motion displays depend on the cortex and on an intact corpus callosum?

A display similar to Fig. 2a subtending 5° was produced on the CRT and presented completely within one hemi-field alone. Two spots were flashed on opposite corners of a square (frame 1) and then replaced by two spots appearing on the two remaining corners²³. The display is perceptually reversible (like a Necker cube)²⁴ and bi-stable in that there are two mutually exclusive percepts, vertical motion and horizontal motion. The two possible percepts (Fig. 2b and c) were explained to each subject (each hemisphere?) and the subject was 'trained' to use the corresponding hand to indicate the direction of perceived movement by simply pointing to an illustration on a card depicting the appropriate direction of movement. Informal testing suggested that each hemisphere could see these alternative percepts quite clearly.

When normal subjects fixate the centre of such a display, there is a slightly greater tendency to see vertical^{25,26} rather than horizontal oscillation, although either percept can be seen voluntarily. As this tendency is usually attributed to delays across the midline imposed by the corpus callosum^{25,26}, we wondered whether commissurotomy patients would always see exclusively vertical oscillation and we tested this conjecture in L.B., N.G. and A.A. To our surprise, we found that when they fixated a small spot in the centre of the display, they could in fact very easily see either organization at will. A perceptual switch could also be encouraged by momentarily occluding portions of the

display; for example, they could switch from vertical to horizontal oscillation by temporarily occluding the lower half of the display and then removing the occluder.

Next, we measured the tendency to see vertical compared with horizontal motion by changing the breadth/height ratio of the ambiguous figure while the subjects fixated an 'X' in the centre of the display. The height of the figure was varied randomly from trial to trial and on each trial the subject's task was simply to indicate whether he/she saw vertical or horizontal motion (inter-trial interval = 10 s; stimulus onset asynchrony = 150 ms; ISI = 0; the breadth was held constant at 5° and a total of 10 different height/breadth ratios were used, ranging from 0.53 to 2.15). For the normal 'control' subjects, vertical and horizontal motion were equally probable when the height/breadth ratio was 1.37 (mean of 40 readings; 20 readings for each subject). For the three commissurotomy patients, the ratios were 1.39, 1.5 and 1.51 for L.B., N.G. and A.A., respectively (mean of 20 readings for each subject; in N.G. we discarded 3 trials on which there were accidental eye movements). These results suggest that commissurotomy patients can see horizontal oscillations across the vertical meridian almost as easily as could normal subjects. Note that when they were seeing horizontal motion across the vertical meridian, the vertical motion signal must be either inhibited or vetoed at some level. The curious implication is that this inhibition can occur across the midline in the absence of forebrain commissures.

Further experiments are in progress to determine whether apparent motion of more abstract stimuli, such as cyclopean targets²⁷, 'subjective contours'^{28,29} and coloured spots at isoluminance³⁰, can also be mediated by the second visual system in these patients. (Such stimuli are normally thought to require cortical processing.) Also, one wonders whether each hemisphere can independently 'will' a perceptual reversal without involving the other hemisphere, a question that might interest theologians.

We thank R. W. Sperry and E. Ebbesen for providing facilities, and C. Hamilton, V. Inada, J. Bogen, F. H. C. Crick, O. Braddick and J. A. Deutsch for stimulating discussions. The experiments were carried out in the laboratory of Dr R. W. Sperry at Caltech, Pasadena, California. V.S.R. was supported by academic senate and biomedical research grants from UCSD. A.C.-G. was supported by the NIH and by a Del E. Webb Fellowship from the California Institute of Technology.

Received 6 August 1985; accepted 23 January 1986.

- Wertheimer, M. *Z. Psychol.* **61**, 161-265 (1912).
- Kolers, P. *Aspects of Motion Perception* (Pergamon, New York, 1972).
- Bogen, J. E. & Vogel, P. J. in *Les Syndromes de Disconnexion Calleuse Chez l'Homme* (eds Michel, F. & Schott, B.) 227-251 (Hôpital Neurologique, Lyon, 1975).
- Bogen, J. E. in *Clinical Neuropsychology* (eds Heilman, K. M. & Valenstein, E.) (Oxford University Press, 1979).
- Sperry, R. W. *Am. Psychologist* **23**, 723-733 (1968).
- Sperry, R. W., Gazzaniga, M. S. & Bogen, J. E. in *Handbook of Clinical Neurology* (eds Vinken, P. J. & Bruyn, G. W.) 273-290 (North-Holland, Amsterdam, 1969).
- Ramachandran, V. S. & Cobb, S. (in preparation).
- Mitchell, D. E. & Blakemore, C. B. *Vision Res.* **10**, 49-54 (1970).
- Choudhury, B. P., Whitteridge, D. & Wilson, M. E. *Q. J. exp. Physiol.* **50**, 214-219 (1965).
- Hubel, D. H. & Wiesel, T. N. *J. Neurophysiol.* **30**, 1561-1573 (1967).
- Zeki, S. M. *Brain Res.* **34**, 63-75 (1970).
- Pettigrew, J., Ramachandran, V. S. & Bravo, H. *Brain. Behav. Evol.* **24**, 64-93 (1984).
- Braddick, O. J. *Vision Res.* **14**, 519-527 (1974).
- Korte, A. *Z. Psychol.* **72**, 194-296 (1915).
- Trevarthen, C. B. & Sperry, R. W. *Brain* **96**, 547-570 (1973).
- Trevarthen, C. B. *Psychol. Forsch.* **31**, 299-337 (1968).
- Schneider, G. E. *Science* **163**, 895-902 (1969).
- Weiszcrautz, L., Warrington, E. K. & Saunders, M. *Brain* **97**, 709-728 (1974).
- Pöppel, E., Held, R. & Frost, D. *Nature* **243**, 295-296 (1973).
- Barlow, H. B., Blakemore, C. & Pettigrew, J. D. *J. Physiol., Lond.* **193**, 327-342 (1967).
- Hubel, D. H. & Wiesel, T. N. *Nature* **225**, 41-42 (1970).
- Ramachandran, V. S., Clarke, P. G. H. & Whitteridge, D. *Nature* **268**, 333-335 (1977).
- Ramachandran, V. S. & Anstis, S. M. *Nature* **304**, 529-531 (1983); *Perception* **14**, 135-143 (1985).
- Necker, L. A. *Phil. Mag.* **1** (3d s.) 329-337 (1832).
- Gengerelli, J. A. *J. exp. Psychol.* **38**, 592-599 (1948).
- Ramachandran, V. S. & Anstis, S. M. *Vision Res.* **23**, 83-85 (1983).
- Julesz, B. *Foundations of Cyclopean Perception* (University of Chicago Press, 1971).
- Ramachandran, V. S., Rao, V. M. & Vidyasagar, T. R. *Vision Res.* **13**, 1399-1401 (1973).
- Ramachandran, V. S. *Perception* **14**, 127-134 (1985).
- Ramachandran, V. S. & Gregory, R. L. *Nature* **275**, 55-56 (1978).

Sampling in spatial vision

Dennis M. Levi & Stanley A. Klein

College of Optometry, University of Houston—University Park,
4901 Calhoun Boulevard, Houston, Texas 77004, USA

The human visual system is capable of making spatial discriminations with extraordinary accuracy. In normal foveal vision, relative position, width or size can be judged with an accuracy much finer than the size or spacing of even the smallest foveal cones. This remarkable accuracy of spatial vision has been termed 'hyper-acuity'¹. Almost a century ago Ewald Hering proposed that the accuracy of Vernier acuity could be accounted for by averaging of discrete samples along the length of the lines comprising the targets²; however, the discovery that Vernier acuity of a few arc seconds could be achieved with dots has rendered the nature and role of sampling in spatial discrimination unclear³. We have been investigating the sampling of spatial information in central and peripheral vision (the periphery) of normal human observers and in observers with strabismic amblyopia. Our results, presented here, show that peripheral vision and central vision of strabismic amblyopes differ qualitatively in their sampling characteristics from those of the normal fovea. Both the periphery and the central visual field of strabismic amblyopes demonstrate marked positional uncertainty which can be reduced by averaging of spatial information from discrete samples.

In the first experiment highly practised observers judged whether a test line presented briefly bisected the interval between two continuously viewed reference lines. The lines were composed of discrete samples (dots), each approximately 1 arc min and separated by inter-sample spaces of variable extent. We varied both the number of dots comprising each line and the inter-space size. The inset in Fig. 1 shows an example of a stimulus in which each line comprised five samples. The vertical separation of the lines was chosen to be optimal for each observer based on preliminary testing (see Table 1, column 2). A signal detection methodology was used to obtain thresholds for this spatial discrimination⁴. Figure 1a shows for two normal observers that foveal thresholds improved only slightly as the number of samples increased from 1 to 5. Additional samples beyond 5 had no further effect on thresholds and with even a single sample, thresholds were smaller than the inter-cone spacing (~ 30 arcs in the fovea).

Figure 1b shows results at 2.5° in the lower visual field; the data differ both quantitatively and qualitatively from those of the fovea. First, the threshold for a single sample is higher by more than a factor of 10. The threshold for a single sample (in the absence of spatial averaging) may provide an estimate of the intrinsic positional uncertainty of the visual system. What is of special interest here is that the threshold for a single sample is much larger than the inter-cone spacing at 2.5° (~ 70 arcs), suggesting that it is the cortical sampling grain rather than the cone mosaic which limits position discrimination in the periphery. Second, in the periphery, adding samples has a strong effect on bisection thresholds. For an 'ideal detector' that is limited by the spatial sampling grain, the positional uncertainty would be reduced in proportion to the square root of the number of independent samples; for the log-log axes of Fig. 1 this would mean a slope of -0.5 . A flatter slope (like that of the fovea) would indicate that sparseness of sampling was not the factor limiting spatial discrimination. A higher slope would indicate an effect on the visibility of the targets by luminance summation. Table 1 gives the parameters for the fits to all the data in Fig. 1 (broken lines). While the results for the fovea are not compatible with a slope of -0.5 , those for the periphery are consistent with a decrease in threshold proportional to $\sqrt{n}$ (n is the number of discrete samples) up to $n = 10$. Beyond this, adding more

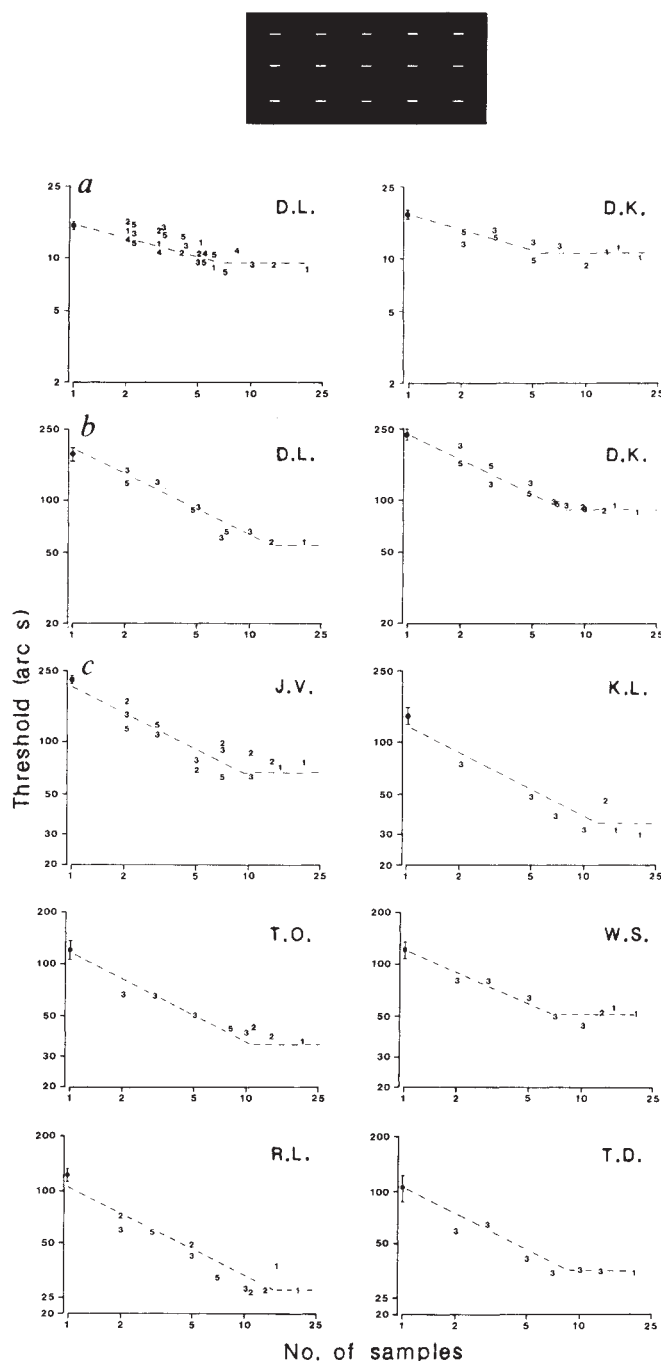


Fig. 1 Top, a schematic diagram of the 3-line bisection stimulus, with each line composed of five samples (that is, dots ~ 1 arc min) each separated by 2 arc min. The observer's task was to judge the vertical position of the briefly flashed test line (middle line) relative to the bisection point. The vertical separation between the reference lines and bisection point was optimized for each panel and is given in Table 1. Bisection thresholds equivalent to the signal detection parameter, $d' = 0.675$ (75% correct), are based on 250-600 trials per point, and are plotted as a function of the number of samples for normal foveal vision (a), peripheral vision (b) and the central field of strabismic amblyopes (c). The numbers in each graph represent the sizes of the inter-space intervals in arc min. Each graph has been fitted with two lines. Thresholds decrease as the number of samples increases to ~ 5 for foveal vision and 10 for peripheral and amblyopic vision. A slope of -0.5 (that is, thresholds proportional to the square root of the number of samples) provides a reasonable fit for the data of the periphery and of the strabismic amblyopes. A shallower slope, -0.3 , provides a good fit to the foveal results (see Table 1).

Table 1 Parameters for the fits to the data in Fig. 1

Condition	Verical separation (min)	Observer	Snellen acuity	Threshold for a single sample (arc s) (y_0)	Asymptotic no. of samples (x_A)	Slope (r)
Normal fovea	3	D.L.	6/4.5	16.4 ± 0.80	7.8 ± 1.7	$-0.29 \pm 0.04^*$
	3	D.K.	6/5	17.4 ± 0.65	5.4 ± 1.3	$-0.30 \pm 0.04^*$
Periphery (2.5°)	12.4	D.L.	6/21	191.6 ± 11.3	11.1 ± 2.0	-0.48 ± 0.04
	12	D.K.	6/22	233.5 ± 10.6	9.5 ± 0.9	-0.43 ± 0.04
Strabismic amblyopia†	10	J.V.	6/24	219.8 ± 25	5.3 ± 0.9	-0.66 ± 0.11
	5	T.D.	6/8	97.7 ± 10.5	7.0 ± 1.4	-0.52 ± 0.08
	6.7	R.L.	6/12	110.8 ± 8.1	10.0 ± 2.6	-0.58 ± 0.09
	12	T.O.	6/10	104.7 ± 8.3	11.6 ± 2.6	-0.43 ± 0.06
	5	K.L.	6/12	114.8 ± 17.5	16.2 ± 7.8	-0.50 ± 0.07
	10	W.S.	6/13.5	117.5 ± 9.1	7.2 ± 1.7	-0.42 ± 0.07

Values were obtained by nonlinear regression of the form $\log y = \log y_0 + r \log x$ for $x < x_A$ and $y = \log y_0 + r \log x_A$ for $x \geq x_A$ where y is the threshold in arc s and x is the number of samples. The SAS statistical package was used. Fitting all the data with a second model, where x is the length of samples plus inter-spaces, gave a significantly worse χ^2 value: 226 as opposed to 172 ($p < 0.025$, F test with 127 d.f.). Errors represent 1 standard error. All observers were carefully refracted and wore appropriate spectacle corrections if necessary. Extensive practice was given before the actual data collection. D.L. is an author. Snellen acuity (75% correct) was derived from crowded Davidson-Eskridge charts (6/6 corresponds to a 60-arc s feature). * Not consistent with a slope of -0.5 at the 0.000001 level, whereas all the other data are consistent with a slope of -0.5 at the 0.05 confidence level. † Each of these observers has a constant unilateral strabismus of early onset and reduced acuity in an otherwise healthy eye.

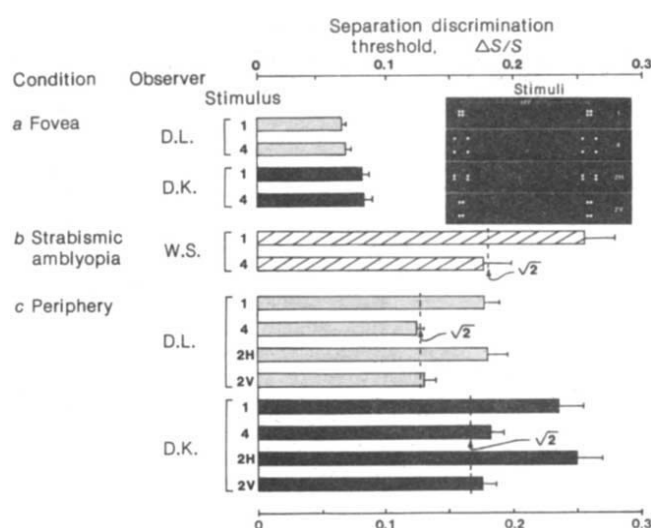
samples has no additional effect on thresholds. The effect of adding samples is not simply explained by luminance summation, as making a single sample seven times brighter in a control experiment did not improve the bisection threshold.

The spatial vision of strabismic amblyopes has previously been compared to that of the normal periphery⁵⁻⁷. The data shown in Fig. 1c for the amblyopic eyes of six strabismic amblyopes bear a remarkable similarity to the results for the normal periphery (Fig. 1b). For each of the strabismic amblyopes, as in the normal periphery, position uncertainty is high with only one sample, and thresholds improve in proportion to the square root of the number of discrete samples. Table 1 also shows Snellen acuity for each observer. For the amblyopic eyes and the periphery, the threshold for one sample is approximately equal to Snellen acuity. In foveal vision, on the other hand, the threshold for one sample is about three times better than Snellen acuity (that is, a hyperacuity). Interestingly, the non-amblyopic eyes of strabismic amblyopes were similar to those of the normal observers. For example, J.V., a highly experienced observer who had the most severe amblyopia, showed a threshold with one sample of 19.0 ± 0.9 arc s and a slope of -0.24 ± 0.1 with his preferred eye.

A second experiment was performed in order to verify our sampling hypothesis and to determine whether information is

sampled only along the length of the lines of the target (that is, orthogonal to the discrimination cue) as originally suggested by Hering². Here, the observers' task was to judge whether the separation between two small boxes was larger or smaller than the standard (12.4 arc min) separation⁸. Each box was actually composed of four tiny dots, illustrated schematically in the inset in Fig. 2 (stimuli 1 and 4). In this experiment, we manipulated the distances between the four dots comprising each of the two boxes. When they were close together (10 arc s) and thus unresolved, they represented a single (bright) sample. When the dots were all separated so that they could be resolved (1.2 arc min for the normal fovea; 2.4 arc min for peripheral and amblyopic vision) they provided four discrete samples. By this device the total luminance of each box stimulus was balanced and was approximately 20 times the observers' detection threshold. Figure 2a shows the separation discrimination threshold for unresolved (one sample) and resolved dots (four samples). For the normal fovea (Fig. 2a) thresholds for the two conditions were identical. For a strabismic amblyopic eye (Fig. 2b) and for the normal periphery (Fig. 2c), thresholds improved noticeably for the resolved dots. We have confirmed these observations on other observers and at different separations. We anticipated that increasing the number of samples from one to four would result in a twofold decrease in threshold (equal to $\sqrt{4}$). However,

Fig. 2 Inset, a schematic diagram of the stimuli for spatial interval discrimination. In an experimental run, the stimulus (two boxes) was presented briefly with one of five closely spaced separations, centred at 12.4 arc min, and the observer judged whether the spatial interval between the two boxes (SEP in inset) was larger or smaller than the implicit standard. Stimulus feedback was given after each trial. Each box actually comprised four tiny bright dots on a dark background. When the dots were closely spaced (10 arc s) and therefore unresolved, there was only one sample per box (stimulus 1). When the dots were separated by 1.2 arc min (fovea) or 2.4 arc min (periphery and strabismic amblyopia) they were seen as four discrete samples (stimulus 4). Separating the dots either only horizontally (2H) or only vertically (2V) resulted in two discrete samples. Threshold, equivalent to $d' = 0.675$, is specified as a fraction of the base separation (12.4 arc min) for unresolved (stimulus 1) and resolved (4) dots for the fovea of two normal observers (a), a strabismic amblyope (b) and peripheral vision of two observers at 2.5° (c). Each threshold is the mean of four counterbalanced runs (125 trials per run). To test which samples were effective in lowering the threshold, the dots were separated either horizontally or vertically. The results show that two samples separated horizontally (2H) were not significantly better than a single sample, whereas two samples separated vertically (2V) were $\sqrt{2}$ better than a single sample in lowering the separation discrimination threshold.



surprisingly, the thresholds were in fact reduced by only $\sim\sqrt{2}$, suggesting that only two of the samples were effective. To test this possibility, we separated the dots either horizontally only or vertically only (stimuli 2H and 2V in the inset), thus providing two samples either in the same direction as the discrimination cue (horizontal separation) or in the orthogonal direction (vertical separation). Horizontal separation gave thresholds identical to that obtained with a single sample whereas vertical separation gave thresholds which were $\sqrt{2}$ better. Thus, the two samples in the direction orthogonal to the discrimination cue each contributed effectively to reducing the threshold. This result is consistent with Hering's hypothesis regarding the averaging of discrete samples and suggests that this process is performed by oriented mechanisms.

Our results show that spatial information is indeed sampled discretely along the length of targets as originally proposed by Hering. In normal foveal vision, this process has a relatively small impact on the accuracy of spatial discriminations, presumably because the fovea has little intrinsic positional uncertainty. Thus foveal thresholds are smaller than a cone diameter. However, in the normal periphery and in the central field of strabismic amblyopes, the addition of spatial samples in the direction orthogonal to the discrimination cue reduces thresholds in proportion to $\sqrt{n}$, as would be expected in an ideal detector with uncorrelated noise at an early stage of visual processing.

In peripheral vision, the high degree of spatial uncertainty with a few samples can be understood on the basis of the anatomy and physiology of the retina and cortex which results in a sparse neural sampling grain⁹. In strabismic amblyopia, we hypothesize that abnormal binocular interactions result in similar neural consequences—that is, a sparse spatial sampling grain as a result of there being insufficient cortical neurones to provide accurate position signals and/or a scrambling of the neural signals⁵. A sparse sampling grain and/or scrambling of neural signals would introduce positional noise which is uncorrelated between stimulus samples in peripheral and strabismic amblyopic vision.

We thank H. Bedell, B. Breitmeyer, R. Harwerth, R. Manny, E. Smith, S. Steinman, G. Westheimer and Y. Yap for helpful comments, and Alvenia Daniels for preparation of the manuscript. This research was supported by research grants R01EY01728 and R01EY04776 from the National Eye Institute.

Received 23 December 1985; accepted 24 January 1986.

1. Westheimer, G. *Invest. Ophthalm. vis. Sci.* **12**, 570–572 (1975).
2. Hering, E. *Ber. Math. phys. Classe Königl. Sachs. Ges. Wiss. Leipzig (Naturwiss. Teil)* **16**–24 (1899).
3. Ludvig, E. *Am. J. Ophthalm.* **36**, 139–142 (1953).
4. Levi, D. M., Klein, S. A. & Aitsebaomo, A. P. *Vision Res.* **25**, 963–977 (1985).
5. Levi, D. M. & Klein, S. A. *Vision Res.* **25**, 979–991 (1985).
6. Braddick, O. *Nature* **298**, 224–225 (1982).
7. Hess, R. & Bradley, A. *Nature* **287**, 463–464 (1980).
8. Westheimer, G. & McKee, S. P. *Vision Res.* **17**, 940–967 (1977).
9. Dow, B. M., Snyder, R. G., Vautin, R. G. & Bauer, R. *Expl. Brain Res.* **44**, 213–228 (1981).

Generation of end-inhibition in the visual cortex via interlaminar connections

Jürgen Bolz & Charles D. Gilbert

Laboratory of Neurobiology, Rockefeller University, New York, New York 10021, USA

To understand the mechanisms by which the receptive field properties of visual cortical cells are generated, one must consider both the thalamic input to the cortex and the intrinsic cortical connections. In the cat striate cortex, layer 4 is the main recipient of input from the lateral geniculate nucleus, yet the cells in that layer possess several receptive field properties that are distinct from the geniculate input, including orientation specificity, binocularity, directionality and end-inhibition, the last of which allows cells to respond to edges of a restricted length^{1–4}. These properties could be generated by connections within the layer, by its input from the claustrum⁵ or by the massive projection that layer 4 receives from layer 6 (refs 6–9). In the present study, we attempted to determine the functional role of the layer 6 to layer 4 projection by reversible inactivation of layer 6 using the inhibitory transmitter γ -aminobutyric acid (GABA). After inactivating layer 6, cells in layer 4 lost end-inhibition. Cells in layer 2+3, which receive their principal input from layer 4, were similarly affected. The elimination of end-inhibition was specific, other receptive field properties, such as direction selectivity or orientation specificity, remaining intact.

Recordings were made in cats maintained on sodium thiopental anaesthesia, paralysed with succinylcholine and artificially respired with 100% oxygen. At the start of an experiment virtually all of the cells in layers 2, 3 and 4 were end-inhibited to some degree. The oxygen helped maintain the proportion of end-inhibited cells, which otherwise tended to decline during the experiment. The animals' electrocardiogram, electroencephalogram, temperature and expired CO₂ concentration were continuously monitored. A small hole was drilled in the skull above the visual cortex, and the dura and pia opened. To

study the role of the layer 6 to 4 pathway, we inactivated layer 6 by injecting GABA and examined the effect of this treatment on the receptive field properties in layers 2 to 4. A similar approach has been used in other systems with inhibitory transmitters or analogues, local anaesthetics or cobalt^{10–12}. The advantage of GABA is that it affects cells and not afferents and the effects are reversible. In initial experiments to determine the area inactivated by GABA injections of various amounts and concentrations, the micropipette was placed about 300 μ m below the layer 5/6 border and a second tungsten electrode, laterally displaced at various distances from the first electrode, was advanced to approximately the same laminar position as the micropipette. The entrance into layer 6 was recognized by the appearance of cells possessing long receptive fields in this layer³. After the penetrations, electrolytic lesions were made to verify the placement of the electrodes histologically and to determine the horizontal displacement between the two electrodes. We found that a 0.1- μ l injection of 10 mM GABA inactivated an area of $\sim 700 \mu$ m in diameter. The inactivation lasted for several minutes and was always reversible.

Having established the conditions for blockade of the activity of layer 6 cells, we recorded from cells in layers 2+3 and 4 while the GABA pipette was in a corresponding topographical position in layer 6. Figure 1 shows response histograms of a simple cell in layer 4. When tested with a stationary flashing bar, the cell's receptive field consisted of a separate 'on' region flanked by two antagonistic 'off' regions. The cell was end-inhibited: its response to a long bar (8°) was reduced by 54% compared with its response to a bar of optimal length (1°). An injection of 0.1 μ l of 10 mM GABA in layer 6 had no effect on the cell's response to the short bar, but greatly increased its response to the long bar, so that end-inhibition was completely eliminated. This blockade was reversible, for 3 min after GABA injection the cell's response was again reduced over 50% by end-inhibition.

Cells in the superficial layers showed similar effects to those seen in layer 4. Figure 2 shows response histograms of a layer 2+3 cell with a complex receptive field, 0.5° $\times$ 1.5° in size, and on-off responses to a stationary flashing bar. The cell responded briskly when a bar of optimal length (0.5°) was moved across its receptive field, but the response was reduced by 58% when

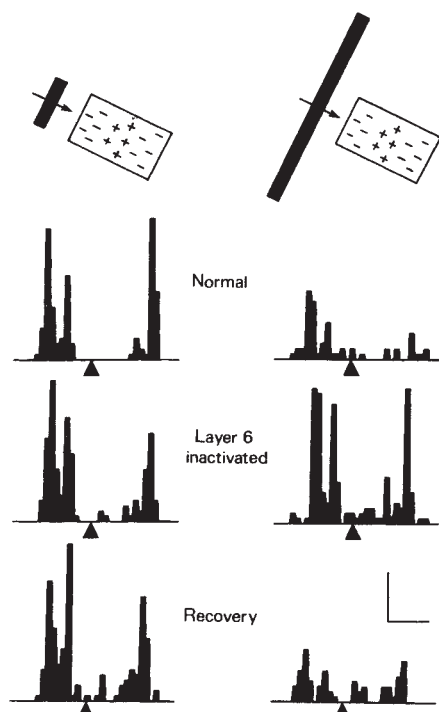


Fig. 1 Effect of inactivation of layer 6 on a simple cell in layer 4. The cell's receptive field was $1^\circ \times 1\frac{1}{2}^\circ$ and had antagonistic on (+) and off (-) subfields when tested with a stationary flashing bar. Layer 6 was inactivated by pressure injection of $0.1 \mu\text{l}$ GABA (10 mM in 0.9% saline) using a glass micropipette with an insulated tungsten electrode glued to the side. Recordings were made either through the tungsten electrode or the micropipette, and were used to determine the duration of blockade. The injection pipette was positioned in layer 6 by advancing it until cells with receptive field properties characteristic of the layer were found³ and then advanced an additional 300 μm . A second tungsten electrode was placed in layer 4, directly over the pipette in layer 6. The positions of the injection pipette and tungsten electrode were later confirmed histologically. Response histograms from 10 stimulus presentations each were made before (top), immediately after (middle) and 3 min after (bottom) GABA injection in layer 6. The triangles under each histogram indicate the point at which the bar reversed direction. The cell was end-inhibited, such that its response to a long bar (8°) was 54% less than its response to a bar of optimum length (1°). This inhibition was eliminated by inactivation of layer 6 (right column), whereas the cell's response to a short bar was unchanged. Calibration: vertical mark, 10 spikes per bin; horizontal mark, 1 s.

the stimulating bar was lengthened to 8° . The cell showed directional preference, responding 2.3 times more to movement in the optimal direction than to movement in the opposite direction. When layer 6 was inactivated by injecting $0.2 \mu\text{l}$ of 10 mM GABA, the cell's response to the short bar was unchanged, but its response to the long bar was strongly increased, so that end-inhibition was eliminated. There was no effect on direction selectivity. Thus, end-inhibition was selectively eliminated, decreasing from 58% before GABA injection to 3% during layer 6 inactivation. Seven minutes after GABA injection the cell had almost completely recovered (to 45% end-inhibition).

For most cells, the reduction of end-inhibition after a GABA injection was readily recognized within the first few responses to a long bar swept across the receptive field. The maximal effect lasted 1–5 min and the cell's response then gradually recovered, being complete within 3–40 min. The effects were reversible after several injections. In one case, 11 injections were made without causing permanent changes in the receptive field properties of the recorded cell and without causing damage visible in Nissl-stained sections.

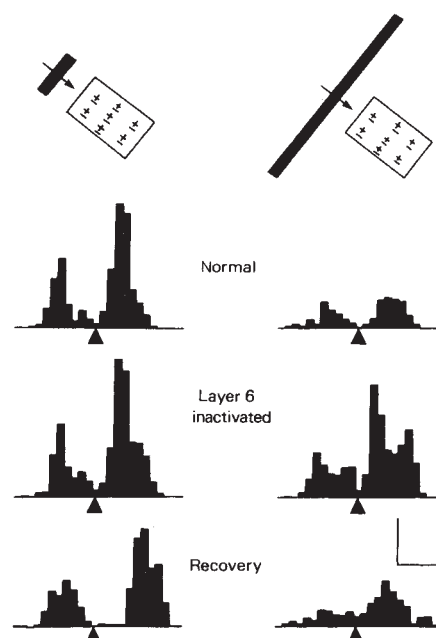


Fig. 2 Effect of $0.2 \mu\text{l}$ GABA (10 mM) injection in layer 6 on a complex cell in layer 2+3. Same procedure as in Fig. 1. The cell's receptive field was $0.5^\circ \times 1.5^\circ$ and gave mixed on (+) and off (-) responses when tested with a stationary flashing bar. When the stimulating bar was 8° long (top right), the cell's response was reduced by 58% of its response to a bar of optimal length (0.5° , top left). During layer 6 inactivation, the cell was only 3% end-inhibited (middle right). Seven minutes after GABA injection the cell was again 45% end-inhibited (bottom right). Inactivation of layer 6 had no effect on the response to the bar with an optimal length of 0.5° (histograms in left column). Note that inactivation of layer 6 had no effect on the cell's direction selectivity. Each histogram was obtained from 10 stimulus presentations. Calibration: vertical mark, 50 spikes per bin; horizontal mark, 2 s.

To determine the spatial distribution of the effect, we compared length response curves before and during inactivation by producing and reversing the effect many times and obtaining histograms before and during inactivation for each of a number of bar lengths. An example is shown in Fig. 3. The cell showed summation until the stimulating bar was increased up to the full length of the excitatory part of the cell's receptive field; further lengthening of the bar inhibited the cell's response until the response was completely suppressed by a bar 4° in length. Inactivation of layer 6 uncovered a vigorous response to long bars, but had no effect on the response to bars equal in length or shorter than the excitatory portion of the receptive field.

We also investigated the effect of layer 6 inactivation on orientation tuning and directionality. The full range of orientations over which we could elicit a response was determined using a hand-held projector, and we determined this range before and during layer 6 inactivation. While layer 6 was inactivated, the orientation tuning to a long bar was sharper than to a short bar, as it was before the blockade. A few cells showed slight increases in width of orientation tuning, although even when end-inhibition was completely eliminated they were still clearly orientation-selective. For the entire population studied directionality was unaffected during layer 6 inactivation (Fig. 2).

The effect of layer 6 inactivation was examined on a sample of 49 end-inhibited cells from 12 cats. We selected only those cells with moderate to strong end-inhibition, and limited the size of the injection to that required for a clear reduction of end-inhibition, but never applied more than $0.3 \mu\text{l}$ in a single injection. The reduction of end-inhibition and the time course of the effect depended on the alignment of the two electrodes and the size of the GABA injection. For example, end-inhibition

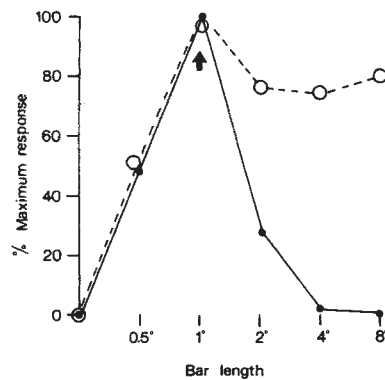


Fig. 3 Length-response curve for a complex cell in layer 2+3 before (continuous line) and after (dashed line) GABA injection in layer 6. The cell showed summation until the stimulating bar extended for the full length of the excitatory portion of the receptive field, as determined by testing its response to a short bar centred at different positions along its axis of orientation (arrow). Further lengthening of the bar reduced the response of the cell until it was completely inhibited by bars of 4° length or longer. Inactivation of layer 6 revealed a vigorous response to long bars, but had no effects on the response to bars equal in length or shorter than the excitatory portion of the receptive field.

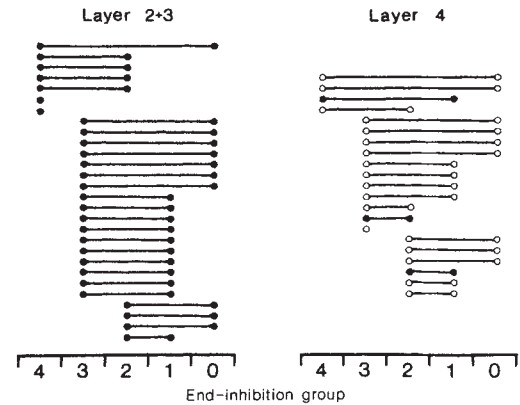


Fig. 4 End-inhibition of cells in layers 2+3 (left) and 4 (right) before and after GABA injection in layer 6. Degree of end-inhibition was divided into five categories. Cells in group 0 had no end-inhibition (0–10%); in group 1, end-inhibition was weak (11–35%); in group 2, it was medium (36–60%); in group 3, it was strong (61–90%); and in group 4, end-inhibition was very strong (91–100%). For all but three cells, inactivation of layer 6 significantly reduced end-inhibition. Closed circles, complex cells; open circles, simple cells.

of the cell in Fig. 2 was reduced from 58% to 40% by 0.1 μ l GABA and abolished by 0.2 μ l. However, the amount of GABA necessary to reduce end-inhibition was independent of the laminar position of the cell, with complex cells in layer 2+3 requiring the same average dosage as simple cells in layer 4. Out of our sample of 49 cells, end-inhibition was significantly decreased for all but 3 of the cells. For the cells where no effect was seen, later histological reconstruction of the positions of the recording and injection electrodes showed that their horizontal displacement was greater than 500 μ m. Figure 4 shows the extent of the change in end-inhibition for our sample population of cells in layers 2+3 and 4.

Our experiments suggest that layer 6 is the source of end-inhibition for cells in layers 4 and 2+3. Many layer 6 pyramidal cells have axon collaterals that project to and ramify within layer 4 (refs 6–8). As mentioned above, these cells have long receptive fields, often requiring relatively long bars for activation and showing increased response as the stimulating bar is lengthened up to 8° or more. Cells in layer 4 and above have small receptive fields, and show progressive inhibition as the stimulating bar is lengthened, up to sizes equivalent to the excitatory portion of layer 6 cell receptive fields. A model for generating end-inhibition from cells with long receptive fields was originally suggested by Hubel and Wiesel¹³. The projection pattern and receptive field properties of layer 6 cells led to the hypothesis that these could be the cells responsible for end-inhibition, either by direct inhibitory action or by making contact with inhibitory interneurons within layer 4 (refs 6, 14).

Layer 6 cells are likely to be excitatory, because they take up and transport aspartate, a putative excitatory transmitter¹⁵, and form asymmetrical synapses with round vesicles, a morphology usually associated with excitatory synapses¹⁶. However, using serial electron microscope reconstructions, McGuire *et al.*¹⁶ demonstrated that many of the processes that are postsynaptic to layer 6 collaterals belong to either smooth or sparsely spinous cells, which are thought to be inhibitory, as they form symmetrical synapses^{17,18} and are GABAergic^{19,20}. A role for GABAergic interneurons in end-inhibition is supported by iontophoresis of the GABA antagonist bicuculline while recording from end-inhibited cells in layer 2+3, though usually the

effects were weak²¹. The fact that the response to short bars was unaffected by layer 6 inactivation is consistent with our model, in that layer 6 cells are frequently not activated by short bars and therefore are unlikely to influence the response of layer 4 cells to short bars. Finally, this hypothesis predicts that some inhibitory interneurons in layer 4 will have long receptive fields. These neurons may have been missed because interneurons are smaller and less common than principal neurons, and it is difficult to determine the receptive field length without appropriate quantitative methods. However, using intracellular recording and dye injection, A. Humphrey (personal communication) recently found a smooth stellate cell in layer 4 possessing a long receptive field similar to those of layer 6 neurones.

Our findings are comparable to those of Sherk and LeVay²², who studied the role of the projection from claustrum to cortex. The claustrum projects to all layers of visual cortex, but most heavily to layer 4 (ref. 5). Cells in the claustrum have similar receptive fields to those in layer 6 (ref. 5), and when the claustrum is lesioned, cortical cells lose some of their end-inhibition²². The effects of inactivating layer 6 seen in the present study were stronger than those caused by lesioning the claustrum. Since layer 6 is the source of input to the claustrum and since this pathway is retinotopically organized²³, inactivating layer 6 could effectively block the claustral input to layer 4 as well as the direct layer 6 to 4 projection, thus removing both sources of end-inhibition with a single injection.

Inactivating layer 6 abolished end-inhibition in layer 2+3 as readily as in layer 4, although layer 6 projects predominantly to layer 4. However, spiny stellate cells in layer 4 have a strong, presumably excitatory projection to layer 2+3 (refs 6, 9). Our results therefore suggest that cells in layer 4 endow layer 2+3 neurones with end-inhibition, an idea which is further supported by the finding that layers 2+3 and 4 have the same proportions of end-inhibited cells³ showing a comparable degree of end-inhibition³. This may explain why bicuculline iontophoresis could not eliminate end-inhibition in the superficial layers, since cells to which the bicuculline was applied may not have generated end-inhibition themselves, but instead may have inherited it from cells distant from the iontophoretic pipette. Our findings provide indirect evidence for an excitatory link between layers 4 and 2+3, and suggest that at least one property is transferred

through that pathway. This conflicts with the finding that layer 2+3 cells appear normal when layer 4 is silenced by inactivation of the A laminae of the geniculate, requiring all receptive field properties to be generated anew in each cortical layer¹¹. Instead, our results are consistent with the hypothesis of Hubel and Wiesel¹ of a sequence of processing from simple cells to complex cells, representing a functional interaction between cortical layers.

In conclusion, we have shown an association between one

component of the intrinsic cortical circuit and a specific receptive field property. Our results demonstrate how a functionally defined class of cortical cells participates in the transformation of geniculate input occurring within the striate cortex.

We thank Torsten Wiesel for his continuous encouragement and support, Larry Katz for his helpful discussions, and Barbara McGuire and Dan T'so for comments. This work was supported by the Max-Planck-Gesellschaft, NIH grant EY04782 and NSF grants BNS8351738 and BNS8318794.

Received 27 September; accepted 12 December 1985.

- Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **160**, 106-154 (1962).
- Dreher, B. *Invest. Ophthalmol.* **11**, 355-356 (1972).
- Gilbert, C. D. *J. Physiol., Lond.* **268**, 391-421 (1977).
- Rose, D. J. *J. Physiol., Lond.* **271**, 1-23 (1977).
- LeVay, S. & Sherk, H. *J. Neurosci.* **1**, 956-980 (1981).
- Gilbert, C. D. & Wiesel, T. N. *Nature* **280**, 120-125 (1979).
- Lund, J. S., Henry, G. H., MacQueen, C. L. & Harvey, A. R. *J. comp. Neurol.* **184**, 599-618 (1979).
- Katz, L. C., Burkhalter, A. & Dreyer, W. J. *Nature* **310**, 498-500 (1984).
- Martin, K. A. C. & Whitteridge, D. J. *Physiol., Lond.* **253**, 463-504 (1985).
- Malpeli, J. G., Schiller, P. H. & Colby, C. L. *J. Neurophysiol.* **46**, 1102-1119 (1981).

- Malpeli, J. G. *J. Neurophysiol.* **49**, 595-610 (1983).
- Newsome, W. T., Wurtz, R. H., Dürsteler & Mikami, A. *Expl Brain Res.* **58**, 392-399 (1985).
- Hubel, D. H. & Wiesel, T. N. *J. Neurophysiol.* **28**, 229-289 (1965).
- Wiesel, T. N. & Gilbert, C. D. *Q. Jl exp. Physiol.* **68**, 525-543 (1983).
- Baughman, R. W. & Gilbert, C. D. *J. Neurosci.* **1**, 427-439 (1981).
- McGuire, B. A., Hornung, J.-P., Gilbert, C. D. & Wiesel, T. N. *J. Neurosci.* **4**, 3021-3033 (1984).
- LeVay, S. *J. comp. Neurol.* **150**, 53-86 (1973).
- Peters, A. & Fairén, A. *J. comp. Neurol.* **181**, 129-172 (1978).
- Ribak, C. E. *J. Neurocytol.* **7**, 461-478 (1978).
- Hamos, J. E., Davis, P. L. & Sterling, P. J. *comp. Neurol.* **217**, 449-457 (1983).
- Sillito, A. M. & Versiani, V. *J. Physiol., Lond.* **273**, 775-790 (1977).
- Sherk, H. & LeVay, S. *J. Neurosci.* **3**, 2121-2127 (1983).
- LeVay, S. & Sherk, H. *J. Neurosci.* **1**, 993-1002 (1981).

Retrovirus insertion inactivates mouse $\alpha 1(I)$ collagen gene by blocking initiation of transcription

Stefan Hartung*, Rudolf Jaenisch*† & Michael Breindl*

* Heinrich-Pette-Institut für experimentelle Virologie und Immunologie an der Universität Hamburg, Martinistrasse 52, 2000 Hamburg 20, FRG

† Whitehead Institute for Biomedical Research and Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, Massachusetts 02142, USA

Mov13 mice carry a single Moloney murine leukaemia virus (M-MuLV) proviral copy in the first intron of the $\alpha 1(I)$ collagen gene. Virus insertion interferes with the synthesis of stable $\alpha 1(I)$ collagen messenger RNA and causes a recessive lethal mutation^{1,2}. The virus insertion has induced changes of the methylation pattern³ as well as the chromatin conformation in the mutated gene. Specifically, a DNase-hypersensitive site which is associated with active transcription of the wild-type collagen gene is not present in the mutant allele⁴. The block of collagen expression could be caused by virus-induced instability of collagen mRNA or by impaired initiation of transcription. To distinguish between these possibilities, we have compared the activity of the $\alpha 1(I)$ collagen gene promoter in cell lines derived from wild-type and Mov13 embryos by nuclear run-on transcription experiments and S₁ mapping of nuclear RNA. We show here that initiation of transcription of the mutant gene is reduced 20-100-fold. This indicates that the virus-induced change of chromatin structure in the promoter region of the mutant gene prevents RNA polymerase from binding to its DNA template. Our results are consistent with the notion that the promoter-associated DNase-hypersensitive site is a prerequisite for rather than a consequence of gene activity.

To determine the transcriptional initiation at the wild-type and Mov13 $\alpha 1(I)$ collagen gene promoter, we performed nuclear run-on transcription experiments with nuclei from cell lines derived from wild-type or homozygous Mov13 embryos¹. Undifferentiated P19 embryonal carcinoma (EC) cells, which do not transcribe $\alpha 1(I)$ collagen mRNA⁴, were used as a control. Nuclei were isolated and RNA was synthesized *in vitro* in the presence of ³²P-UTP and hybridized to a Southern blot of restriction fragments which were derived from a 14-kilobase (kb) clone containing the 5' end of the $\alpha 1(I)$ collagen gene¹.

Restriction fragments of a rat α -tubulin complementary DNA clone⁵ served as internal standard. The result of a typical experiment is shown in Fig. 1a. *In vitro* synthesized RNA from all cell lines hybridized with comparable intensity to the α -tubulin DNA fragment. In contrast, only RNA from wild-type, but not from Mov13 or P19, cells hybridized to the $\alpha 1(I)$ collagen restriction fragments. (Fragments 1B and 2B, which did hybridize, contain repetitive sequences located 5' of the $\alpha 1(I)$ collagen gene, see map, Fig. 1c). The intensity of hybridization of the individual fragments was roughly proportional to their size, indicating that sequences derived from different sections of the gene were transcribed *in vitro* with identical frequency. Moreover, hybridization of the 1.1-kb *Bst*EII/*Xba*I fragment 1D, which contains only intron sequences (Fig. 1c), shows that the *in vitro* elongated RNAs used as hybridization probes were primary transcripts and not processed RNA.

The 0.5-kb *Bst*EII/*Bgl*II fragment 2C contains almost exclusively sequences derived from the first exon of the $\alpha 1(I)$ collagen gene which in Mov13 DNA is located 5' of the viral integration site. The intensity of hybridization of this fragment is therefore a measure of initiation of $\alpha 1(I)$ collagen mRNA transcription in wild-type as well as in Mov13 cells. Figure 1a shows that fragment 2C hybridized only to RNA from wild-type but not from Mov13 cells. Similarly, no signal was seen with RNA from P19 cells. Even after extended exposure of the autoradiogram (Fig. 1b) the 0.5-kb fragment 2C did not give a visible signal with RNA from Mov13 cells. This shows that in Mov13 cells the absence of $\alpha 1(I)$ collagen mRNA is due, at least in part, to a reduction of the number of RNA polymerase molecules initiating RNA transcription at the $\alpha 1(I)$ collagen promoter rather than a premature termination or aberrant splicing of transcripts. Densitometer scans of autoradiograms from several independent experiments (not shown) allowed us to estimate that initiation of $\alpha 1(I)$ collagen gene transcription in Mov13 cells is at least 20-fold lower than in wild-type cells. However, because of the small size of fragment 2C and the corresponding weak signal produced on the autoradiogram, this is a minimal estimate indicating the limits of detection by this assay rather than the actual extent of transcriptional inhibition.

In vitro elongated RNA from Mov13 cells did hybridize to the 4.9-kb fragment 1F (Fig. 1b; a faint hybridization of fragment 1F with P19 RNA was also visible on the original autoradiogram but is not shown). To determine whether Mov13 and P19 cells show residual $\alpha 1(I)$ collagen promoter activity which was undetectable by the nuclear run-on transcription assay but could account for the hybridization of Mov13 transcripts to fragment

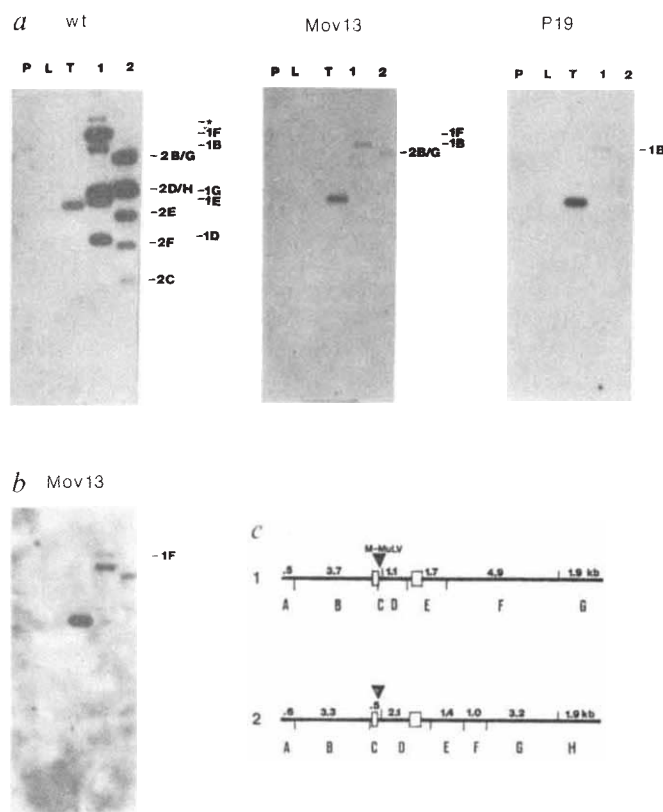


Fig. 1 Hybridization of *in vitro* elongated RNA to restriction fragments of the mouse $\alpha 1(I)$ collagen gene. Nuclei were prepared from cell lines derived from wild-type (wt) or homozygous Mov13 embryos¹ and from the EC cell line P19. Nuclear RNA was elongated in the presence of ³²P-UTP, purified and hybridized to nitrocellulose filters containing restriction fragments of a genomic clone of the mouse $\alpha 1(I)$ collagen gene (lanes 1, 2 and ref. 2), a rat α -tubulin cDNA clone (lane T and ref. 5), and pBR322 (P) or λ (L) control DNA. The filters were exposed for 1 day (a) or 8 days (b). A partial restriction map of the mouse $\alpha 1(I)$ collagen gene² is shown in c. The first two exons of the gene are shown as open boxes, the M-MuLV integration site in Mov13 DNA as an arrowhead. The restriction sites of the enzymes *BstEII*, *XbaI* and *BglII* are shown which were used to produce the DNA blots shown in a. The uppermost band marked with an asterisk in lane 1 is due to a partial digest.

Methods. Isolation of nuclei, run-on transcription, purification of labelled RNA, and hybridization were performed according to published procedures¹⁴⁻¹⁶, with minor modifications. After exposure to X-ray film, the filters were rehybridized with a nick-translated mouse $\alpha 1(I)$ collagen DNA probe to confirm that all the blots used contained identical amounts of DNA (not shown).

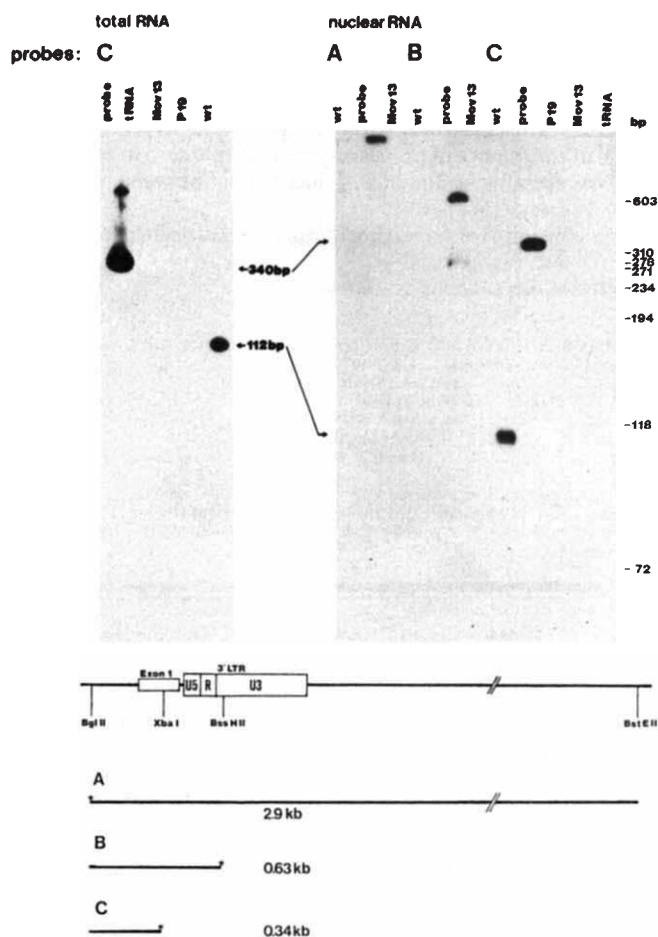


Fig. 2 Nuclease *S*₁-mapping analysis of total and nuclear RNA from wild-type (wt), homozygous Mov13 and P19 cells. The RNAs analysed with probes A, B or C, respectively (compare the map and the text), are indicated at the top of the autoradiograms.

Methods. Cells were homogenized in hypotonic buffer⁴ in the presence of 0.03–0.05% Triton X-100 and nuclei isolated by centrifugation and extracted with phenol-chloroform. RNA was separated from DNA by CsCl gradient centrifugation. Samples of 15 μ g of nuclear or 25 μ g of total RNA were hybridized to 10–20 ng of each probe, digested with *S*₁ nuclease and analysed by polyacrylamide gel electrophoresis².

1F, we performed nuclease *S*₁ mapping experiments with nuclear RNA from wild-type, Mov13 and P19 cells. These RNA preparations showed a ~100-fold enrichment of nuclear RNA sequences, as estimated from the yield of RNA recovered in this fraction relative to total cellular RNA and from an electrophoretic analysis of 45S ribosomal RNA precursor molecules in the nuclear RNA preparations (not shown). Because of the presence of considerable amounts of this latter RNA species, we also expected an enrichment of primary mRNA transcripts in the nuclear RNA fraction and therefore an increased sensitivity of the nuclease *S*₁ mapping assay.

The probes used in these assays and the results are shown in Fig. 2. Besides a 0.34-kb *BglII/XbaI* fragment (probe C) which has previously been used for *S*₁ nuclease mapping of $\alpha 1(I)$ collagen mRNA^{2,4}, we have also used a 0.63-kb *BglII/BssHII* fragment (probe B) prepared from a genomic clone of Mov13

DNA⁶. This fragment, which contains the first exon of the $\alpha 1(I)$ collagen gene and sequences derived from the 3' long terminal repeat (LTR) of the M-MuLV provirus integrated in Mov13 DNA was used to detect any transcripts initiating at the collagen promoter and continuing into viral sequences. The 2.9-kb *BglII/BstEII* fragment (probe A) was labelled at the opposite strand and was included to detect possible transcripts initiating at the viral promoter in the 3' LTR and extending into cellular sequences (that is, anti-sense RNA complementary to the first exon and 5'-flanking sequences of the $\alpha 1(I)$ collagen gene). Probe C detected the expected 112-base-pair (bp) protected fragment in both total and nuclear RNA from wild-type but not from Mov13 or P19 cells (Fig. 2). Accordingly, probe B did not detect any transcripts in Mov13 nuclear RNA which would have resulted in a protected fragment of 400 bp. This shows that there is no residual $\alpha 1(I)$ collagen promoter activity in Mov13 cells

which might have escaped detection by the run-on transcription assay. The sensitivity of the S_1 mapping assays was estimated by counting the radioactivity in the corresponding gel slices (not shown). In nuclear RNA from Mov13 and P19 cells we would have been able to detect a 100-fold lower concentration of $\alpha 1(I)$ collagen transcripts as compared with wild-type cells. Assuming that the nuclear RNA used contained predominantly primary transcripts, we can conclude that initiation of transcription at the $\alpha 1(I)$ collagen promoter in Mov13 cells is reduced by at least 100-fold relative to wild-type cells and is comparable to P19 cells which normally do not transcribe $\alpha 1(I)$ collagen mRNA. The hybridization of restriction fragment 1F with *in vitro* elongated RNA from Mov13 cells (Fig. 1b) can therefore not be due to transcripts initiated at the $\alpha 1(I)$ collagen promoter. It may be the result of activation of cryptic promoter sequences within the $\alpha 1(I)$ collagen gene due to the provirus insertion, or cross-hybridization of this section of the gene with type III collagen mRNA or with other unknown repetitive sequences. The presence of a member of the mouse *B1* family of middle repetitive sequences in the $\alpha 1(I)$ collagen gene has in fact recently been shown⁷.

Probe A did not detect any transcripts initiating in the proviral 3' LTR and extending into cellular sequences (Fig. 2). This was not unexpected because indirect evidence indicates that this proviral copy in Mov13 cells is transcriptionally inactive⁴, and because intact retroviral proviruses usually do not use their 3' LTR for initiation of transcription⁸.

Our results show that the provirus-induced change of chromatin structure of the $\alpha 1(I)$ collagen gene in Mov13 cells inhibits initiation of transcription, probably by preventing RNA polymerase and/or other transcription factors from binding to the DNA template. The most striking difference in chromatin structure of the wild-type and mutant gene is the absence of the transcription-associated DNase-hypersensitive site in the latter⁴. DNase-hypersensitive sites within eukaryotic genes are thought to be regions where transcription factors bind to regulatory DNA sequences⁹⁻¹². They may have an important role in developmental gene activation¹³. Because the virus insertion in Mov13 cells has not changed the DNA sequence containing the presumptive binding site but exerts its effect over a distance of ~500 bp, the binding of such factors must be controlled by structural features of the chromatin rather than DNA sequence alone. Our results therefore suggest that the promoter-associated hypersensitive site is a prerequisite for rather than a consequence of gene activity. The virus seems to cause the mutation by long-range and indirect effects which may prevent correct activation of the $\alpha 1(I)$ collagen gene by interfering with the developmentally regulated induction of the hypersensitive site.

We thank K. Harbers for the gift of plasmids and C. Dony and U. Chen-Betdeken for advice with the run-on transcriptions. This work was supported by the Deutsche Forschungsgemeinschaft and by grants HD-19015 from NIH and P01-CA38497 from the NCI. The Heinrich-Pette-Institut is supported by Freie und Hansestadt Hamburg and by Bundesministerium für Jugend, Familie und Gesundheit.

Received 9 December 1985; accepted 19 January 1986.

1. Schnieke, A., Harbers, K. & Jaenisch, R. *Nature* **304**, 315-320 (1983).
2. Harbers, K., Kuehn, M., Delius, H. & Jaenisch, R. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1504-1508 (1984).
3. Jähner, D. & Jaenisch, R. *Nature* **315**, 594-597 (1985).
4. Breindl, M., Harbers, K. & Jaenisch, R. *Cell* **38**, 9-16 (1984).
5. Lemischka, T. R., Farmer, S., Racaniello, V. R. & Sharp, P. A. *J. molec. Biol.* **151**, 101-120 (1981).
6. Jaenisch, R. *et al.* *Cell* **32**, 209-216 (1983).
7. Monson, J. B., Friedman, J. & McCarthy, B. *Molec. cell. Biol.* **2**, 1362-1371 (1982).
8. Cullen, B. R., Lomedico, P. T. & Ju, G. *Nature* **307**, 241-245 (1984).
9. Elgin, S. C. R. *Cell* **27**, 413-415 (1981).
10. Wu, C. *Nature* **309**, 229-234 (1984).
11. Emerson, B. M., Lewis, C. & Felsenfeld, G. *Cell* **41**, 21-30 (1985).
12. Dynan, W. S. & Tjian, B. *Nature* **316**, 774-778 (1985).
13. Weintraub, H. *Cell* **42**, 705-711 (1985).
14. Groudine, M., Peretz, M. & Weintraub, H. *Molec. cell. Biol.* **1**, 281-288 (1981).
15. Lacy, E., Hardison, R. C., Quon, D. & Maniatis, T. *Cell* **18**, 1273-1283 (1979).
16. Dony, C., Kessel, M. & Gruss, P. *Nature* **317**, 636-639 (1985).

The *trans*-activator gene of HTLV-III is essential for virus replication

Amanda G. Fisher*, Mark B. Feinberg*, Steven F. Josephs*, Mary E. Harper*, Lisa M. Marselle*, Gregory Reyes†, Matthew A. Gonda‡, Anna Aldovini*, Christine Debouk§, Robert C. Gallo* & Flossie Wong-Staal*

* Laboratory of Tumor Cell Biology, Developmental Therapeutics Program, National Cancer Institute, Bethesda, Maryland 20205, USA
† Gene Labs Incorporated, 871 Industrial road, San Carlos, California 94070, USA

‡ Laboratory of Cell and Molecular Structure, National Cancer Institute, Frederick Cancer Research Facility, Frederick, Maryland 21701, USA

§ Department of Molecular Genetics, Smith Kline & French Laboratories, 709 Swedeland Rd, Swedeland, Pennsylvania 19479, USA

Studies of the genomic structure of human T-lymphotropic virus type III (HTLV-III) and related viruses, implicated as the causal agent of acquired immune deficiency syndrome (AIDS), have identified a sixth open reading frame in addition to the five previously known within the genome (*gag*, *pol*, *sor*, *env* and 3'*orf*)¹⁻⁴. This gene, called *tat*-III, lies between the *sor* and *env* genes and is able to mediate activation, in a *trans* configuration, of the genes linked to HTLV-III long terminal repeat (LTR) sequences⁵⁻⁸. We now present evidence that the product of *tat*-III is an absolute requirement for virus expression. We show that derivatives of a biologically competent molecular clone of HTLV-III⁹, in which the *tat*-III gene is deleted or the normal splicing abrogated, failed to produce or expressed unusually low levels of virus, respectively, when transfected into T-cell cultures. The capacity of these *tat*-III-defective genomes was transiently restored by co-transfection of a plasmid clone containing a functional *tat*-III gene or by introducing the TAT-III protein itself. As HTLV-III and related viruses are the presumed causal agents of AIDS and associated conditions¹⁰⁻¹², the observation that *tat*-III is critical for HTLV-III replication has important clinical implications, and suggests that specific inhibition of the activity of *tat*-III could be a novel and effective therapeutic approach to the treatment of AIDS.

Transcriptional activation of viral genes by viral regulatory elements acting in *cis* and *trans* configurations was first described for DNA tumour viruses¹³⁻¹⁸, and *trans*-activator genes have since been found in retroviruses of man¹⁹⁻²¹, cattle^{22,23} and sheep²⁴. Knowledge that *trans*-acting factors can specifically augment the expression of viral genes has led to intense speculation that *trans*-activator genes may have crucial roles in the biology of these viruses.

The *tat*-III gene, as characterized by functional mapping studies, consists of three exons^{6,7}. The first exon is non-coding and extends 287 base pairs from the 5' LTR. The second exon, located between the *sor* (for short open reading frame) and *env* (envelope) genes, encodes 72 amino acids and is crucial for *trans*-activation. The third exon encodes only 14 amino acids and is located within the *env* gene in a different reading frame. To investigate whether *trans*-activation is an essential feature of HTLV-III replication, a panel of HTLV-III plasmid clones containing modifications in the *tat*-III gene were generated (Fig. 1). These constructs were derived from a plasmid clone pHXB2D which contains full-length HTLV-III provirus and produces infectious HTLV-III virions and cytopathic effects when transfected into normal human T lymphocytes⁹. Clone pHXB2gpt, which contains an *Xba*I/*Hpa*I viral DNA fragment from λ HXB2D inserted into plasmid vector pSP65gpt, was used to generate genomic constructs lacking either the major coding region (pHXB2 Δ Sal-Sst) or the splice acceptor sequences

Fig. 1 Construction of HTLV-III plasmids and analysis of CAT activity. The precise nature of the deletions shown was confirmed by nucleotide sequence analysis. The nucleotide positions shown correspond to those published for BH10 (ref. 1). Closed boxes depict LTR sequences, cross-hatched boxes depict the *Escherichia coli* xanthine-guanine phosphoribosyl transferase gene (*xgpt*)²⁹ and open triangles show regions deleted from the HTLV-III genome. Open squares shown in plasmids pCV-1 and pCV-3 (ref. 6) represent adenovirus middle late promoter (MLP) sequences³⁰.

Methods. An 11.5-kb *Xba*I/*Hpa*I fragment containing full-length HTLV-III proviral DNA was excised from the phage clone λ HXB2-D (ref. 31) and inserted into the *Bam*HI and *Eco*RI sites of the plasmid pSP65gpt. Clone pHXB2gpt contains the viral insert in the same transcriptional orientation as the *E. coli* *xgpt* gene. pHXB2gpt was used for the construction of pHXB2 Δ Sal-Sst and pHXB2 Δ Sal-RI plasmids. Construct pHXB2 Δ Sal-Sst is deleted of nucleotides 5,367–5,580 and hence lacks the entire first coding exon of *tat*-III (exon 2), including the initiating methionine. It was generated from pHXB2gpt by a *Sst*I partial restriction digest followed by digestion to completion with *Sal*I, end repair with T4 DNA polymerase and re-closure with T4 DNA ligase. Plasmid pHXB2 Δ Sal-RI is deleted of nucleotides 5,323–5,367 and hence lacks the splice acceptor sequence thought to be preferentially used to generate mature *tat*-III mRNA⁶, while the entire *tat*-III coding sequence is unaffected. The clone was constructed by an *Eco*RI partial restriction of pHXB2gpt, followed by complete digestion with *Sal*I, repair with the Klenow fragment of *E. coli* DNA polymerase and re-closure with T4 ligase. Plasmids pCV-1 and pCV-3 contain the complete cDNA of the spliced *tat*-III and 3'orf mRNA, respectively⁶. The ability of plasmid clones to *trans*-activate genes linked to HTLV-III LTR sequences was assessed as follows: 10 μ g of DNA from the recombinant plasmid pCD12CAT (containing the bacterial gene for CAT under the regulation of the HTLV-III LTR²⁶) was mixed with 10 μ g of DNA from each test plasmid individually and transfected into 1.1×10^7 H9 lymphoid cells. The transfections were performed by incubating the cells in RPMI 1640 media containing 250 μ g ml⁻¹ DEAE-dextran, 50 mM Tris-HCl pH 7.3 for 1 h at 37 °C. Cells were washed with complete media (RPMI 1640, 10% fetal calf serum (FCS) and 50 μ g ml⁻¹ gentamicin) and maintained for 48 h at 37 °C in complete media. The cells were collected, washed in phosphate-buffered saline and resuspended in 1 ml of 0.04 M Tris-HCl pH 7.4, 0.15 M NaCl and 0.001 M EDTA, transferred to Eppendorf tubes and pelleted. Cells were resuspended in 0.25 M Tris-HCl pH 7.8 and lysed by freeze-thawing three times. After pelleting cell debris (5 min in Beckman microfuge), supernatants were transferred to new Eppendorf tubes and heated to 60 °C for 10 min. CAT activity was determined on 20- μ l aliquots incubated with ¹⁴C-chloramphenicol and acetyl CoA as previously described²³. Chloramphenicol and acetylated metabolites were separated by ascending TLC and visualized by autoradiography. Radiolabelled chloramphenicol and derivatives were cut from the plates and quantitated by liquid scintillation counting. CAT activity was determined as counts per min (c.p.m.) of acetylated metabolites of chloramphenicol expressed as a percentage of the total c.p.m. The results shown are the means and standard deviation of three independent co-transfections for each test plasmid. Transfection of RSV-CAT³², SVOCAT³² and pCD12CAT alone gave values of 48.0 ± 6.9 , 0.31 ± 0.036 and 1.41 ± 0.13 , respectively. RSV-CAT thus served as a positive control for each experiment, while SVOCAT, which lacks a functional promoter, is a negative control. Levels of CAT activity significantly above the values for pCD12CAT alone were indicative of *trans*-activating activity.

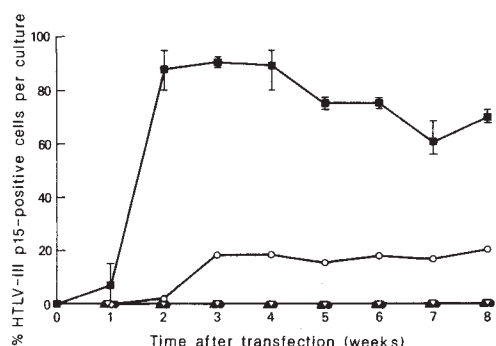
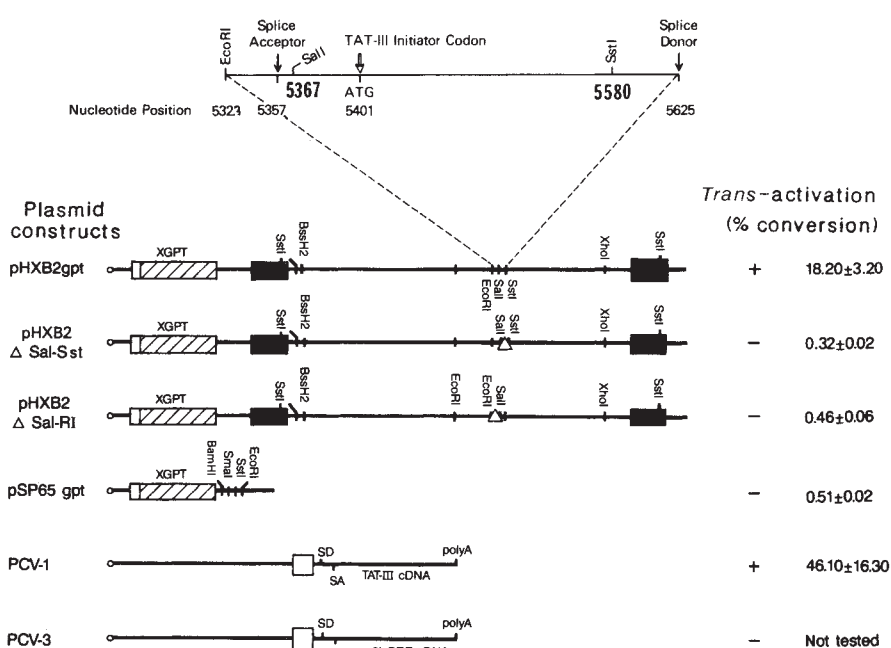


Fig. 2 Comparison of virus expression in H9 cultures following transfection with HTLV-III plasmid derivatives. The percentage of cells expressing HTLV-III p15 antigen was assessed by immunofluorescence using the monoclonal antibody BT2 and standard protocols³³. The results shown are the mean value and range obtained following three independent transfections with plasmids pHXB2gpt (■), pHXB2 Δ Sal-Sst (▲), pSP65gpt (▽) and pHXB2 Δ Sal-RI (cultures 1 (●), 2 and 3 (○)).

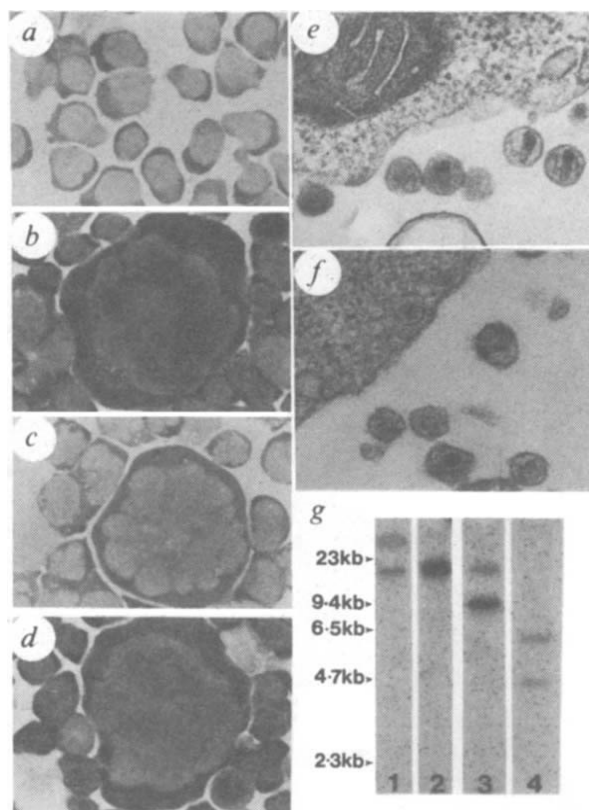
Methods. H9 cells (2.5×10^6) were transfected with 10^{10} bacterial protoplasts using a protoplast fusion technique⁹. The cells were maintained in culture at a density of 5×10^5 to 2×10^6 per ml for 8 weeks in RPMI 1640 media containing 20% fetal calf serum/antibiotics, and HTLV-III virus production assayed by electron microscopy 1, 2, 4, 6 and 8 weeks after transfection.

(pHXB2 Δ Sal-RI) of *tat*-III. In pHXB2 Δ Sal-Sst, nucleotides 5,367–5,580 (from the *Sal*I/*Sst*I sites) have been deleted, removing most of the *tat*-III coding sequences of the second exon together with sequences upstream of the *tat*-III initiation codon (position 5,401) (Fig. 1). Plasmid pHXB2 Δ Sal-RI has lost nucleotides 5,323–5,367 (from the *Sal*I/*Eco*RI sites) and hence lacks the splice acceptor (position 5,357) thought to be preferentially used to generate mature *tat*-III messenger RNA⁶, while the entire coding sequence remains unaffected.

The extent to which these deleted genomes are capable of *trans*-activating genes linked to the HTLV-III LTRs was assessed in a series of DEAE-mediated co-transfection experiments²⁵ using pCD12CAT as an indicator²⁶. Chloramphenicol acetyltransferase (CAT) activity was determined 48 h after transfection as the amount of cytoplasmic ¹⁴C-labelled chloramphenicol converted to acetylated metabolites, expressed as a percentage of the total. Transfection of H9 cells with pHXB2gpt + pCD12CAT significantly augmented CAT activity (18.2%) as compared with transfection using pCD12CAT alone (0.03%) or pCD12CAT + pSP65gpt (0.5%) (Fig. 1), consistent with pHXB2gpt containing a biologically functional *tat*-III gene²⁶. Co-transfection of pCD12CAT with pHXB2 Δ Sal-Sst or pHXB2 Δ Sal-RI gave values of 0.3% and 0.4%, respectively. Thus, in standard assays, *trans*-activation was not detected using genomes that lacked either the second exon coding sequences or the splice acceptor site of the *tat*-III gene. These results are consistent with reports of the location of *trans*-activator gene of HTLV-III and its normal splicing^{6,7}.

To assess the capacity of *tat*-III-defective genomes for productive virus replication, constructs pHXB2 Δ Sal-Sst and

Fig. 3 Demonstration of multinucleated cells, virus particles and HTLV-III DNA sequences in cultures transfected with HTLV-III plasmid derivatives. Three to five days after transfection, H9 cells were removed from culture, cytocentrifuge preparations made and stained with Wright-Giemsa stain to reveal nuclear morphology. *a-d*, Samples from cultures transfected with pSP65gpt (*a*), pXHB2gpt (*b*), pXHB2ΔSal-Sst (*c*) and pXHB2ΔSal-RI (*d*). Multinucleated cells at frequencies of >1 to 5% of total were observed in samples *b*, *c* and *d*. *e, f*, Electron micrographs of virions produced 2 weeks after transfection with pXHB2gpt and pXHB2ΔSal-RI, respectively. In the case of pXHB2gpt, the virions had typical HTLV-III morphology, were abundant and budding particles readily observed. In the case of pXHB2ΔSal-RI, both characteristic HTLV-III virions and particles with abnormal morphology (irregular shape and cores) were evident and very few particles were found. Budding virus was not evident in these cultures 1–4 weeks after transfection but was occasionally demonstrated thereafter. *g*, Southern blots of DNA prepared from H9 cells 2 weeks after transfection with pXHB2ΔSal-Sst (lanes 1–3) or pXHB2gpt (lane 4) DNA of high M_r was isolated using standard protocols³⁴. DNA (5 μg) was digested separately with the enzymes *Xba*I (lane 1), *Bam*HI (lane 2) or *Sst*I (lanes 3, 4) electrophoresed in 0.8% agarose gels, blotted and hybridized to probe prepared from nick translation of the *Sst*I/*Sst*I fragment of the phage clone λBHIO³¹. Nitrocellulose acetate filters were washed in 0.5 × SSC, 0.1% SDS at 65 °C for 2–4 h before exposure. Digestion of H9/pXHB2ΔSal-Sst DNA with *Xba*I (an enzyme that does not cut the viral genome but has a single site in the plasmid) yields a single band at 17 kb, corresponding to linearized plasmid in these cells, and a high- M_r smear (>23 kb), suggesting integration of the transfected DNA. Digestion with *Bam*HI (an enzyme that cuts at a single site in the HTLV-III sequences of the plasmid) reveals a single 17-kb band, consistent with the presence of plasmid DNA in transfected cells. *Sst*I digestion of DNA from pXHB2ΔSal-Sst-transfected cultures shows an intense band of 9.5 kb and a less intense band of 17 kb. The former corresponds to the *Sst*I/*Sst*I viral insert which is predicted following the construction of pXHB2ΔSal-Sst from pXHB2gpt. The latter probably represents residual partially digested plasmid. *Sst*I digestion of H9/pXHB2gpt DNA (lane 4) shows two bands at 5.8 and 3.7 kb, representing the fragments predicted from digestion of unintegrated proviral DNA with this enzyme.



pXHB2ΔSal-RI, pXHB2gpt and pSP65gpt, were introduced into H9 cells by protoplast fusion. This technique has been used by our own and other laboratories to transfect and stably express genes in lymphoid cells^{9,27}. Figure 2 summarizes time-course experiments in which the frequency of cells expressing HTLV-III p15 (*gag*-related protein) was assessed following protoplast fusion. One week after transfection with pXHB2gpt, 1–15% of H9 cells expressed HTLV-III p15 and extracellular and budding virions were evident. The number of HTLV-III-expressing cells increased to 80–90% within the first 2 weeks and was maintained at a high level throughout the experiment. In contrast, no virus- or p15-expressing cells were detected after transfection with pXHB2ΔSal-Sst. These results indicate that *tat*-III is crucial for HTLV-III production. Southern blots (see Fig. 3g) prepared from these cells and probed for HTLV-III sequences showed the presence of linearized unintegrated plasmid DNA (a 17-kilobase (kb) band, lanes 1, 2) and integrated HTLV-III sequences (a high relative molecular mass (M_r) smear >23 kb, lane 1). Hence, the failure of H9/pXHB2ΔSal-Sst cultures to produce virus was not due to failure to introduce this DNA into cells. Notably, digestions with enzymes that do not cut in the HTLV-III sequence (*Xba*I, Fig. 3g, lane 1) or cut at a single site in the plasmid (*Bam*HI, lane 2) failed to reveal a 9.5-kb band, corresponding to unintegrated HTLV-III DNA in pXHB2ΔSal-Sst-transfected cells. The absence of unintegrated viral DNA may be an indication of lack of virus replication.

Transfection of clone pXHB2ΔSal-RI resulted in virus expression in two of three cases. In these, low-level expression (<2% of cells) of p15 was demonstrated 2 weeks after transfection. Electron micrographs revealed small numbers of exclusively extracellular particles, many of which had unusual morphology compared with typical virions (see Fig. 3f and e, respectively). These aberrant particles may represent true defective particles or simply degenerating virions. The proportion of cells expressing p15 increased to ~30% by week 3 and reached a plateau at 10–20% in successive weeks.

These results have two implications. First, the molecular clone pXHB2ΔSal-RI is capable of generating HTLV-III virus, but at

markedly reduced levels compared with the parental pXHB2gpt plasmid (we estimate that cultures transfected with the former contained 10 times fewer particles). Hence, we suggest that this clone, deprived of the normal splice acceptor used to generate *tat*-III, is able to use alternative splice acceptor sites in the 5' region of the genome to generate mature *tat*-III mRNA. Alternatively, we can postulate that *tat* is being read off a polycistronic message which includes *sor-tat-3'orf*. Complementary DNA clones corresponding to these sequences have been identified and shown to *trans*-activate in the CAT assay (S. K. Arya *et al.*, in preparation). An apparent inconsistency with these hypotheses is the failure to demonstrate *trans*-activation with plasmid pXHB2ΔSal-RI. To clarify this point, we performed extended CAT assays on the virus-positive and -negative pXHB2ΔSal-RI-transfected cultures. Cultures 1 (virus negative) and 2 (virus positive) (Fig. 2), assayed 8 weeks after transfection, gave conversions of ¹⁴C-chloramphenicol of $0.534 \pm 0.068\%$ and $52.47 \pm 11.07\%$, respectively, in 13-h assays. The latter value is significantly greater than the conversion rates seen in uninfected cultures (0.5%), consistent with low levels of *trans*-activation in virus-positive, pXHB2ΔSal-RI-transfected H9 cultures. Second, the frequency of HTLV-III p15-positive cells in pXHB2ΔSal-RI-transfected cultures did not increase to approach 100%. Using an *in situ* technique capable of detecting infrequent cells expressing as few as 20–60 HTLV-III RNA transcripts per cell²⁸, we determined that 0% and 25% of cultures 1 and 2 (week 8) and 20% of culture 3 (week 4) expressed viral RNA. These results are comparable to immunofluorescence values and support the contention that virus generated from pXHB2ΔSal-RI is either defective or produced in too low amounts to spread efficiently through the culture. As the splice acceptor deleted in pXHB2ΔSal-RI may be used to generate *tat*-III and 3' *orf* mRNA, the reduced infectivity might be attributable to compromising either or both of these genes.

Figure 3c, d demonstrates the capacity of *tat*-III-defective constructs transiently (3–5 days after protoplast fusion) to induce H9 cell multinucleation at frequencies exceeding 1% (1–5%) and suggests that shortly after transfection with

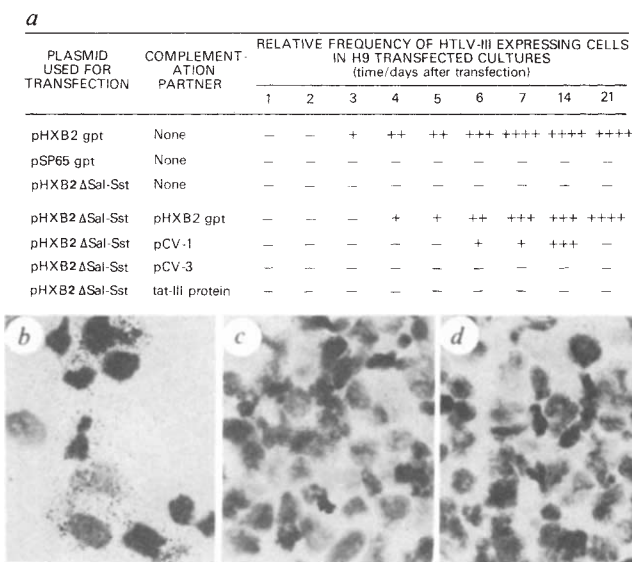


Fig. 4 Complementation of *tat* activity in H9 cells transfected with HTLV-III plasmids deleted in the *trans*-activator gene. **a**, Time course of relative frequency of HTLV-III-expressing cells in H9-transfected cultures. Expression of viral *gag* protein p15 was assessed on a daily basis and the proportions of expressing cells are indicated on an increasing scale where — = none detected, + = few expressing cells, ++ = 0.1–1.0% positive cells, +++ = 1.0–10% positive cells and ++++ = 10–99% positive cells. **b–d**, Autoradiographs of transfected cells labelled by *in situ* hybridization for HTLV-III RNA. The number of grains (intensity of labelling per cell) indicates the relative abundance of HTLV-III RNA. **b**, **c**, H9 cell cultures 16 days after transfection with pHXB2ΔSal-Sst. In the case of **b**, 48 h before *in situ* assay these cultures were transfected with protoplasts containing TAT-III protein. Comparison with parallel cultures not transfected with TAT-III protein (**c**) shows that expression is clearly elevated in the former, 30% and 1% of cells in cultures, corresponding respectively to **b** and **c**, being judged as positive (>20 grains per cell). H9 cells were also transfected with pHXB2ΔSal-Sst and 14 days later transfected with protoplasts containing TAT-III protein. In this case (**d**), we were unable to detect HTLV-III RNA 48 h later.

Methods. **a**, Using the protoplast fusion technique as described previously⁹, 10¹⁰ bacterial protoplasts containing plasmids pHXB2ΔSal-Sst, pSP65gpt or pHXB2gpt were separately transfected into 3 × 10⁶ H9 cells. In addition, 5 × 10⁹ protoplasts containing pHXB2ΔSal-Sst were separately combined with equivalent amounts of pCV-1, pCV-3 and pHXB2gpt or bacterial protoplasts containing TAT-III protein, and the mixed preparations fused with 3 × 10⁶ H9 cells. Bacterial protoplasts containing TAT-III protein were prepared as follows. *E. coli* strain AR120 carrying the plasmid pOTS-*tat*-III (ref. 35) were grown exponentially and induced to synthesize TAT-III protein by the addition of 60 μg ml⁻¹ nalidixic acid³⁶, for 5 h at 37 °C. 10¹⁰ bacteria were collected, washed, treated with lysozyme to produce protoplasts and fused with H9 cells. The efficiency of TAT-III production following nalidixic acid treatment was checked by subjecting bacterial lysates to SDS-polyacrylamide gel electrophoresis and visualization of the separated proteins by Coomassie brilliant blue staining (data not shown). We estimate that the TAT-III protein accounts for 2–5% of the total cellular protein in the induced bacteria. **b–d**, Cells were cytocentrifuged on to microscope slides, fixed and hybridized *in situ* according to Harper *et al.*²⁸. The ³⁵S-labelled RNA probe was synthesized by transcription of clone pBH10-R3²⁸ and was specific for the 3' portion (1–2 kb) of the HTLV-III genome.

pHXB2ΔSal-Sst or pHXB2ΔSal-R1, small numbers of cells transcribe HTLV-III envelope genes. This observation can in part be explained by reports that HTLV-III LTRs can act as efficient promoters in lymphoid cells^{5,26}, in the absence of *tat*-III.

Complementation experiments were performed to determine whether the failure of TAT-defective clones to generate HTLV-III efficiently could be compensated by the provision of a functional *tat*-III gene or the TAT-III protein itself (see Fig. 4). H9 cells transfected with pHXB2gpt alone or in combination with pHXB2ΔSal-Sst expressed HTLV-III p15 and virus particles 3–4 days after transfection, indicating that the simultaneous fusion of protoplasts containing different plasmids does not alter the efficiency of DNA transfection *per se*. HTLV-III-expressing cells were not detected following transfection of pHXB2ΔSal-Sst alone or in combination with pCV-3. However,

following co-transfection of pHXB2ΔSal-Sst and pCV-1, HTLV-III p15-positive cells were apparent 6 (<0.01%) to 14 (2.8%) days after fusion and HTLV-III RNA-expressing cells confirmed in parallel *in situ* studies. These results show that the capacity for virus expression can be restored to a genome deprived of *tat*-III by the provision of an unlinked, functional *tat*-III gene. No virus-expressing cells were seen in the same cultures 21 days after transfection. This transient expression infers that complementation was achieved *in trans* and that decreased expression was coincident with reduced numbers of cells which stably retained both pHXB2ΔSal-Sst and pCV-1. Analysis of DNA prepared from these cells on day 14 showed that the proviral DNA corresponds exactly to that predicted from clone pHXB2ΔSal-Sst (data not shown), and excludes the generation of a competent virus by recombination.

The introduction of bacterially derived TAT-III protein into H9/pHXB2ΔSal-Sst cultures did not result in virus expression, whereas its addition to H9/pHXB2ΔSal-Sst cultures caused a marked increase in virus expression within 48 h; approximately 1% of cells expressed HTLV-III RNA before the introduction of *tat*-III (Fig. 4c), compared with 30%, including heavily labelled (>1,000 grains per cell) multinucleated cells, 2 days after its introduction (Fig. 4b). Thus, a single introduction of TAT-III protein appears insufficient to restore virus production to a genome devoid of *tat*-III coding sequences, but sufficient to boost virus production in a genome containing a functionally compromised *tat*-III gene. This discrepancy may reflect a short biological half-life of the TAT-III protein. The observation that *tat*-III is essential for HTLV-III replication parallels a report that *tat*-II (otherwise called the *x* or *x-lor* gene of HTLV-II) mediates *trans*-activation of viral genes and is essential for replication of HTLV-II²¹. These data support the notion that *trans*-activation control systems are fundamental to the biology of this newly identified group of retroviruses. In addition, as *tat* activity seems to be essential for virus production, the development of inhibitors that specifically block the activity of these *trans*-acting factors provides a new approach to the treatment of AIDS patients and individuals infected with HTLV-III.

We thank Lynn Simek, Laura Stingl, Mary Lipsky and Marty White for their help in these studies and Dr S. K. Arya for providing the cDNA clones pCV-1 and pCV-3. This work was supported in part by grants to A.G.F. from the Foundation for Advanced Education in the Sciences Inc. and the Lady Tata Memorial Research Fund. M.B.F. is a predoctoral fellow in the Cancer Biology Programme of Stanford University.

Received 8 November 1985; accepted 10 February 1986.

1. Ratner, L. *et al.* *Nature* **313**, 277–284 (1985).
2. Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S. & Alizon, M. *Cell* **40**, 9–18 (1985).
3. Sanchez-Pescador, R. *et al.* *Science* **227**, 484–492 (1985).
4. Muesing, M. A. *et al.* *Nature* **313**, 450–458 (1985).
5. Sodroski, J. G. *et al.* *Science* **227**, 171–173 (1985).
6. Arya, S. K., Guo, C., Josephs, S. F. & Wong-Staal, F. *Science* **229**, 69–73 (1985).
7. Sodroski, J. G., Patarca, R., Wong-Staal, F., Rosen, C. A. & Haseltine, W. A. *Science* **229**, 74–77 (1985).
8. Rosen, C. A., Sodroski, J. G. & Haseltine, W. A. *Cell* **41**, 813–823 (1985).
9. Fisher, A. G., Collalti, E., Ratner, L., Gallo, R. C. & Wong-Staal, F. *Nature* **316**, 262–265 (1985).
10. Gallo, R. C. *et al.* *Science* **224**, 500–503 (1984).
11. Barre-Sinoussi, F. *et al.* *Science* **220**, 868–871 (1983).
12. Levy, J. A. *et al.* *Science* **225**, 69–72 (1984).
13. Nevins, J. R., Ginsberg, H. W., Blanchard, J. M., Wilson, M. C. & Darnell, J. E. *J. Virol.* **32**, 727–733 (1979).
14. Jones, N. & Shenk, T. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3665–3669 (1979).
15. Khoury, G. & May, E. *J. Virol.* **23**, 167–176 (1977).
16. Brady, J. N., Bolen, J. B., Radonovic, M., Slazman, N. & Khoury, G. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2040–2044 (1984).
17. Benoist, C. & Chambon, P. *Nature* **290**, 304–310 (1981).
18. Nevins, J. R. *Cell* **26**, 213–220 (1980).
19. Sodroski, J. G., Rosen, C. A. & Haseltine, W. A. *Science* **225**, 381–385 (1984).
20. Sodroski, J. G., Rosen, C. A., Goh, W. C. & Haseltine, W. A. *Science* **228**, 1430–1432 (1985).
21. Chen, I. S. Y. *et al.* *Science* **229**, 54–57 (1985).
22. Rosen, C. A., Sodroski, J. G., Kettman, R., Burny, A. & Haseltine, W. A. *Science* **227**, 320–322 (1985).
23. Derse, D., Caradonna, S. J. & Casey, J. W. *Science* **227**, 317–320 (1985).
24. Hess, J. L., Clements, J. & Narayan, O. *Science* **229**, 482–485 (1985).
25. Queen, C. & Baltimore, D. *Cell* **33**, 741–748 (1983).
26. Seigel, L. *et al.* *Virology* **148**, 226–231 (1986).
27. Yoakum, G. H. *Biotechniques* **25**, 24–30 (1984).

28. Harper, M. E., Marselle, L. M., Gallo, R. C. & Wong-Staal, F. *Proc. natn. Acad. Sci. U.S.A.* **83**, 772-776 (1986).
29. Muligan, R. C. & Berg, P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2072-2076 (1981).
30. Kaufman, J. R. & Sharp, P. A. *Molec. cell. Biol.* **2**, 1304-1319 (1982).
31. Shaw, G. M. *et al. Science* **226**, 1165-1171 (1984).
32. Gorman, C. M., Merlino, G. T., Willingham, M. C., Pastan, I. & Howard, B. H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6777-6781 (1982).
33. Veronese di Marzo, F. *et al. Virology Proc. natn. Acad. Sci. U.S.A.* **82**, 5199-5202 (1985).
34. Maniatis, T., Fritsch, E. F. & Sambrook, J. in *Molecular Cloning: A Laboratory Manual*, 23-233 (Cold Spring Harbor Laboratory, New York, 1982).
35. Aldovini, A. *et al. Proc. natn. Acad. Sci. U.S.A.* (submitted).
36. Mott, J. E., Grant, R. A., Ho, Y. S. & Plat, T. *Proc. natn. Acad. Sci. U.S.A.* **82**, 98-92 (1985).

A model of synthetase/transfer RNA interaction as deduced by protein engineering

Hugues Bedouelle* & Greg Winter

Laboratory of Molecular Biology, Medical Research Council, Hills Road, Cambridge CB2 2QH, UK

The recognition of transfer-RNA by their cognate aminoacyl-tRNA synthetases is the crucial step in the translation of the genetic code. In order to construct a structural model of the complex between the tyrosyl-tRNA synthetase (TyrTS) from *Bacillus stearothermophilus* and tRNA^{Tyr}, 40 basic residues at the surface of the TyrTS dimer have been mutated by site-directed mutagenesis and heterodimers created *in vitro* by recombining subunits derived from different mutants. As reported here a cluster of basic residues (Arg 207-Lys 208) in the N-terminal domain of one TyrTS subunit interacts with the acceptor stem of tRNA^{Tyr} and two separated clusters of basic residues (Arg 368-Arg 371; Arg 407-Arg 408-Lys 410-Lys 411) in the C-terminal domain of the other subunit interact with the anticodon arm. The TyrTS would thus clamp the tRNA in a fixed orientation. The precise alignment of the flexible ... ACCA 3' end of the tRNA for attack on the tyrosyl adenylate is made by contacts closer to the catalytic groups of the enzyme, such as with Lys 151.

TyrTS catalyses the aminoacylation of tRNA^{Tyr} in a two-stage reaction. The tyrosine is first activated with ATP to form tyrosyl adenylate and pyrophosphate, then the adenylate is attacked by the 3'-terminal ribose of the tRNA to form Tyr-tRNA^{Tyr} and AMP. TyrTS is a dimer which shows 'half-of-the-sites' reactivity, forming one tyrosyl adenylate and binding tightly one tyrosine and one tRNA^{Tyr} per dimer in solution. The two subunits are related by symmetry through a 2-fold axis. Each subunit has an

N-terminal domain (residues 1-319), which makes all the interactions with the tyrosyl adenylate and forms the subunit interface, and a C-terminal domain (residues 320-419), which is disordered in the crystal (reviewed in ref. 1). By creating a truncated TyrTS at the level of the gene, Waye *et al.*² have shown that the N-terminal domain of TyrTS catalyses the formation of tyrosyl adenylate with unchanged k_{cat} and K_M but does not charge and does not bind tRNA^{Tyr}; this shows that the C-terminal domain of TyrTS is essential for tRNA binding.

To identify residues of TyrTS that interact with tRNA^{Tyr}, we chose the following strategy: (1) We assumed that basic residues of the synthetase could form salt bridges with the phosphates of the tRNA backbone or hydrogen bonds with the nucleotide bases. (2) Because the *B. stearothermophilus* and *Escherichia coli* TyrTS have homologous sequences and similar properties^{3,4}, we considered mainly the conserved basic residues. (3) We changed the arginine and histidine residues to glutamine, and lysine to asparagine; such changes remove the charge but not the hydrophilic character of the residue. We therefore mutated 40 basic residues of the *B. stearothermophilus* TyrTS by oligonucleotide-directed mutagenesis of the encoding gene (*tyrS*) (see Fig. 1 legend).

To test the overall activity of the mutant synthetases, we devised an *in vivo* genetic complementation assay. In this assay, the *B. stearothermophilus tyrS* gene is carried by and expressed from a recombinant M13 phage⁵. The host is HB2111, an *E. coli* strain which harbours a thermosensitive mutation in its own *tyrS* gene, which is an essential gene. The HB2111 cells can grow at 42 °C, the non-permissive temperature, only if they are infected by a phage which directs the production of an active *B. stearothermophilus* TyrTS. Most of the 40 mutant phages could complement HB2111 and were therefore eliminated. However, 13 mutants were either unable to complement—KN82, RQ86, KN151, KN208, KN230, KN233, RQ368, RQ407, KN410 and KN411—or did so weakly—RQ207, RQ371 and RQ408 (Fig. 1).

The 13 TyrTS mutants identified from the complementation assay and the wild-type TyrTS were purified from phage-infected cells. All the mutant enzymes were able to form enzyme-bound tyrosyl adenylate, albeit slowly for mutants RQ86 and KN233 (half life, $t_{1/2}$ = 8 and 27 min respectively, compared with 2 s for the wild-type enzyme)⁶. The pyrophosphate exchange assay showed that the mutant and wild-type enzymes had similar activities (4.6 s^{-1} at 2 mM ATP, 50 μM Tyr and 2 mM pyrophosphate) except for mutants KN82, RQ86, KN230 and KN233 ($<0.23 \text{ s}^{-1}$) (see Table 1). Thus, although these four mutants are able to form tyrosyl adenylate and are presumably correctly folded, there is nevertheless a lesion in the activation step. Additional experiments are needed to determine whether the

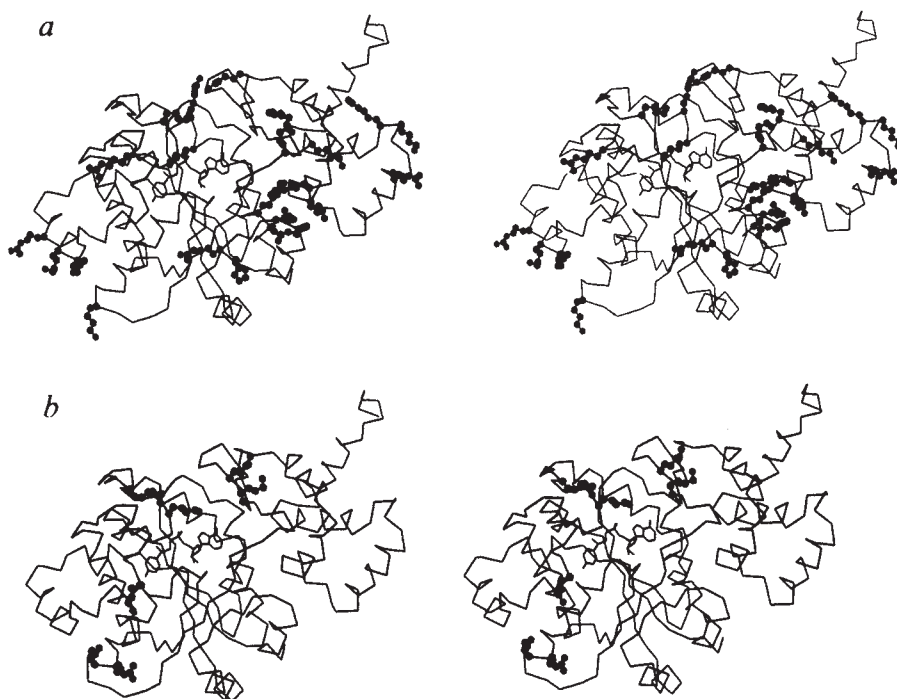
* Present address: UPMTG, Institut Pasteur, 28, Rue du Docteur Roux, 75724 Paris, Cedex 15, France.

Table 1 Kinetic parameters for tRNA^{Tyr} charging

Mutant	k_{cat} ($\text{s}^{-1} \times 10^3$)	K_M (μM)	k_{cat}/K_M ($\text{s}^{-1} \text{ M}^{-1} \times 10^{-3}$)	ΔG_{app} (kcal mol^{-1})	Complementation
Wild type	450	1.4	315	0	+
KN151	3	1.2	2.3	2.9	-
RQ207	ND	>100	19.4	1.7	+/-
KN208	ND	>28	11.3	2.0	-
RQ368	ND	>100	2.4	2.9	-
RQ371	ND	>100	17.1	1.7	+/-
RQ407	ND	>100	11.9	1.9	-
RQ408	ND	>100	16.5	1.8	+/-
KN410	ND	>100	9.3	2.1	-
KN411	ND	>100	12.8	1.9	-

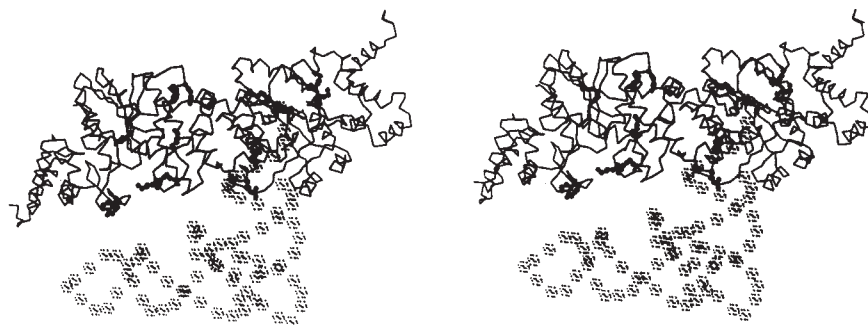
$\Delta G_{app} = -RT \ln(k_{cat}/K_M)_{mut}/(k_{cat}/K_M)_{wt}$ (mut, mutant; wt, wild type)¹¹. For genetic complementation assay, see Fig. 1 legend. ND, not determined. Purification of the wild-type and mutant enzymes from phage-infected cells⁷, active-site titration, pyrophosphate exchange and tRNA charging assays¹¹ were done as described elsewhere. We obtained different values for the wild-type TyrTS with pure *E. coli* tRNA^{Tyr} (Boehringer: 1,000 pmol of tyrosine incorporation per A_{260} unit) and crude *E. coli* tRNAs (K_M = 1.43 versus 2.57 μM and k_{cat} = 0.45 versus 1.17 s^{-1}). The rate of charging by the wild-type enzyme gave a straight line in the Eadie-Hofstee plot (see ref. 12) only for pure tRNA^{Tyr} concentrations less than 6 μM and was inhibited for higher values. We therefore used pure *E. coli* tRNA^{Tyr} for the mutant enzymes, in the range 0.375-6 μM .

Fig. 1 Stereo view of the TyrTS backbone with mutated side chains. Forty basic residues—lysine (K), arginine (R) or histidine (H)—were mutated to asparagine (N) or glutamine (Q). Twenty-seven mutants gave full growth (+) in the genetic complementation assay while 13 mutants gave residual (+/–) or no growth (–): RQ10+, RQ56+, RQ57+, HQ63+, RQ64+, KN82–, KN83+, RQ86–, RQ100+, KN102+, RQ137+, KN141+, HQ142+, KN151–, RQ157+, RQ207+/–, KN208–, KN210+, KA225+, KN230–, KN233–, KN245+, RQ265+, KN268+, RQ286+, KN291+, RQ292+, KN304+, HQ307+, KN367+, RQ368–, RQ371+/–, RQ385+, RQ398+, RQ402+, RQ407–, RQ408+/–, KN410–, KN411–, RQ417+, wild type+. The N-terminal domain (residues 1–319) is represented with bound tyrosyl adenylate. *a*, 22 residues that, when mutated, score as +; *b*, 7 residues that score as +/– or –. From bottom to top of the figure: R207, K208, K151, K82 and R86 on the left, K230 and 233 on the right.



Methods. Primer design: In general, we used codons AAC for Asn and CAG for Gln. CAA was used in the mutants RQ64, RQ292 and RQ402 to improve discrimination during the screening of the mutants. In general, the oligonucleotides contained seven pairing residues on each side of the mismatched base(s) and were 15- or 16-mers. The oligonucleotides for mutants RQ137 and KN268 contained an additional pairing residue on the 3' or 5' side of the mismatch, respectively, to provide a G-C pair at this end of the primer. The oligonucleotide for mutant KA225 had the sequence 5' TC CGC TGC CGT CAC AAG 3'; KA225 was constructed by P. Carter (unpublished). *Oligonucleotide synthesis:* The oligonucleotides were synthesized by Biosearch SAM1 DNA synthesizer using phosphotriester chemistry on a controlled-pore glass support, and were purified by either HPLC on an ion-exchange resin (Partisil 10SAX) or gel electrophoresis using ethidium bromide staining¹⁴. *Mutagenesis technique:* The mutagenesis was done on derivatives of the original M13mp93tyrS vector⁵ that carry an 'up' mutation in the *tyrS* promoter (M. M. Y. Waye and G. W., in preparation) and, in some cases, an amber mutation in M13 gene IV¹⁵. The technique used was "priming all the way round/sucrose gradient" for 9 mutants and "double priming with amber selection" for 31 mutants¹⁴. The mutant phages were screened by hybridization with the mutagenic oligonucleotides¹⁴ followed by dideoxy sequencing using a family of synthetic primers¹⁶. Mutagenic oligonucleotides were also used as sequencing primers. *Complementation assay:* We constructed HB2111 (*tyrS*(Ts) *recA* *srl::tn10 argG thi lac/F' lac*⁺) from the *E. coli* K-12 strain 565cN (ref. 17) and used it as host. About 10⁷ cells of an exponential culture of HB2111, at 30 °C in complete medium, were spread on a pre-warmed minimal glucose plate containing tetracycline (15 µg ml⁻¹) and arginine. Drops (10 µl) of the phage suspensions (~5 × 10¹¹ plaque-forming units (PFU) ml⁻¹) were spotted onto the lawn and the plate was incubated at 42.5–43 °C. Colonies were scored after 48 h. The mutant HN45, which does not activate tyrosine¹, and the truncated TyrTS² were both negative in the assay.

Fig. 2 Docking of tRNA and synthetase. The figure shows the backbone of the TyrTS dimer with tyrosyl adenylate bound to the right subunit and with the side chains of residues R207, K208, K151, K82, R86, K230 and K233 (for the identification of these residues, see Fig. 1 legend). Only the N-terminal domain (residues 1–319) of each subunit is represented. The structure of tRNA^{Phe} is indicated by the van der Waals' spheres of its phosphorous atoms. The flexible 3'-terminus of the tRNA (...ACCA 3') was docked up to the tyrosyl adenylate binding site and the 3'-terminal adenosine could be fitted into the active site so as to attack the tyrosyl adenylate at the scissile bond; however, several rotations of the phosphodiester backbone were required to avoid bad contacts.



wild-type residues at these four positions also interact with the tRNA; in the crystallographic structure, they lie on the rim of the tyrosyl adenylate binding site (Fig. 1).

Table 1 shows the kinetic results for the charging of tRNA by the nine mutants that exchange pyrophosphate at a normal rate. Eight of the nine mutant enzymes have altered K_M values for tRNA in the aminoacylation reaction, which indicates that most of the contacts between the synthetase and the tRNA are involved in the initial complex formation. The KN151 mutant has a K_M for tRNA which is identical to that of the wild-type enzyme and a k_{cat} which is reduced by 150-fold; this indicates that the wild-type residue, Lys 151, is not involved in the initial binding between the synthetase and the tRNA and that TyrTS

and tRNA^{Tyr} may make an additional contact, involving Lys 151, in the transition state.

Previously, we have shown that it is possible to form and purify heterodimers between truncated (lacking residues 319–417) and full-length TyrTS enzymes by mixing the two parental homodimers in equimolar amounts, denaturing the mixture with urea, then renaturing it by electrophoresis through a non-denaturing polyacrylamide gel. After renaturation, the two-refolded homodimers and the heterodimer are roughly in the proportion 1:1:2, as theoretically expected. Detailed kinetic studies of purified heterodimers carrying the Asn 45 mutation (which prevents tyrosine activation) in either the full-length or the truncated subunit, have shown that one tRNA molecule

Table 2 Charging activity of TyrTS heterodimers

Mixture	Aminoacylation rate ($s^{-1} \times 10^3$)	
	a	b
tr+KN82	270	59
tr+RQ86	277	4
tr+KN151	196	2
tr+RQ207	304	29
tr+KN208	328	25
tr+KN230	297	6
tr+KN233	250	9
tr+RQ368	3	2
tr+RQ371	32	22
tr+RQ407	16	11
tr+RQ408	31	23
tr+KN410	14	11
tr+KN411	20	13
KN151+RQ368	162	4

The truncated TyrTS (tr)² and each of the 13 mutants were mixed in equal molarities (according to active-site titration) before (a) or after (b) the denaturation/renaturation process. A similar experiment was performed with the KN151 and RQ368 mutants. The rates are calculated per input mole of one of the two parental homodimers; calculated in this way, they can reach 100% of the rate of the re-folded wild-type enzyme if there is random reassociation of the monomers. In case a, half of the monomers are theoretically engaged in heterodimer formation and therefore the contribution of the re-folded homodimers should be half the value in b. The rates for the re-folded truncated and wild-type enzymes were 0 and $421 \times 10^{-3} s^{-1}$ respectively. The rate of charging by the active TyrTS heterodimers (46–75% that of reconstituted wild-type enzyme) implies that the tRNA affects the formation or the hydrolysis of tyrosyl adenylate. Each heterodimer has two potential sites for tyrosine activation and a single site for tRNA binding. However, due to half-of-the-sites reactivity, only one molecule of tyrosyl adenylate is formed per heterodimer. If tyrosyl adenylate is formed on the truncated subunit, it can be attacked by the tRNA (rate $0.45 s^{-1}$; see Table 1) and the enzyme can recycle, whereas if adenylate is formed on the other subunit, it cannot be transferred to the tRNA and also prevents further tyrosine activation at the truncated subunit. In the latter case, the enzyme is blocked until the tyrosyl adenylate hydrolyses (rate $0.45 \times 10^{-3} s^{-1}$; ref. 13). If tyrosyl adenylate could form at random at either subunit, the heterodimers would quickly be blocked (50% of the remaining active heterodimers each cycle). Thus, we suspect that either the tyrosyl adenylate forms preferentially on the truncated subunit of the heterodimers (directed by the lie of the tRNA), or the binding of tRNA greatly stimulates hydrolysis of blocking tyrosyl adenylate. The experiments were done in 100 mM Tris-HCl, 44 mM Tris pH 7.78, 10 mM MgCl₂, 10 mM 2-mercaptoethanol and 0.1 mM phenyl methyl sulphonyl fluoride. For denaturation, the starting solutions (20 μ l) were 10.4 μ M in each enzyme; 1 μ l bovine serum albumin (BSA, 20 mg ml⁻¹) was added, then urea until saturation (~25 mg; volume 40 μ l, urea 10 M, BSA 0.5 mg ml⁻¹ final concentration). For renaturation, the urea solutions were diluted using nine steps of 1 M each (15 min at room temperature for each step), keeping BSA at 0.5 mg ml⁻¹. Aliquots (2.25 μ l) of the mixtures were assayed in 100 μ l of 10 mM Mg-ATP, 20 μ M ¹⁴C-Tyr (514 mCi mmol⁻¹), 5 mg ml⁻¹ crude *E. coli* tRNA (380 pmol of tyrosine incorporation per mg). The presence of BSA during the denaturation/renaturation process was necessary to recover activity. 73% of the activity of the wild-type TyrTS could be recovered using stepwise dilution of urea whereas only 7% was recovered after a one-step dilution. The addition of urea to the tRNA charging assay resulted in a progressive increase of the rate (2.5-fold maximum at 0.16 M urea).

interacts with both TyrTS subunits⁷. These studies also show that the 3'-terminal adenosine of the tRNA attacks the tyrosyl adenylate in the truncated subunit of the heterodimer.

Here, we have formed heterodimers between the truncated TyrTS and each of the 13 mutants described above by stepwise dilution of the urea. As controls, we also mixed the two parental homodimers after the denaturation/renaturation process. We assayed the final mixtures for charging of tRNA with ¹⁴C-tyrosine (Table 2). The seven mutants carrying a lesion in the N-terminal domain formed active heterodimers with the trun-

cated TyrTS. Similarly, an active heterodimer could be made between the mutant enzymes KN151 and RQ368. Thus, a functional tRNA interaction site can be reconstituted by assembling two inactive monomers, one of which has a wild-type N-terminal domain but lacks the C-terminal domain (or has a lesion therein) and the other having a mutated N-terminal domain but a wild-type C-terminal domain⁷. The results show that tRNA^{Tyr} interacts with the residues Lys 151, Arg 207 and Lys 208 (which correspond to mutated residues) in the truncated subunit of the heterodimer and therefore, taking into account our previous results⁷, that it interacts with both tyrosyl adenylate and residues Lys 151, Arg 207 and Lys 208 in the same subunit of the dimer. The six mutants with a lesion in the C-terminal domain did not form active heterodimers with the truncated TyrTS; this result was expected because the C-terminal domain of TyrTS is necessary for tRNA charging².

In the active heterodimers, one of the constitutive monomers lacks the C-terminal domain. Therefore, in the wild-type TyrTS, only one of the two symmetrical copies of the C-terminal domain is essential for tRNA interaction. Similarly, only one of the two monomers has a wild-type N-terminal domain. Therefore, only one of the two symmetrical copies of residues Lys 151, Arg 207 or Lys 208 is essential for interaction with tRNA. These results show that TyrTS can function with only one productive tRNA binding site per dimer.

Model-building studies⁴ indicate that tRNA^{Tyr} may fold into a structure which is similar to the known structure of yeast tRNA^{Phe} (ref. 8). We have attempted a docking of TyrTS⁹ with tRNA^{Phe} using the computer graphics program FRODO¹⁰. The 3'-hydroxyl of tRNA^{Phe} was held near the tyrosyl adenylate binding site and the tRNA rotated around this fixed point so that the acceptor arm interacted with the tyrosyl adenylate and with residues Lys 151, Arg 207 and Lys 208 on the same subunit. With the path of the tRNA fixed as above, we examined the interaction of the anticodon arm with the second (nonproductive) subunit; this subunit interacts with the tRNA through its disordered C-terminal domain. In the crystallographic map, this domain may correspond to an island of electron density which is located in part on top of the α -helical domain and of the tyrosyl adenylate binding pocket. It is unlikely that the N-terminal domain of the nonproductive subunit makes important contacts to the tRNA through basic residues, as the heterodimers constructed between the truncated TyrTS and the non-complementing mutants at positions 82, 86, 151, 207, 208, 230 and 233 were all active in tRNA charging (Table 2). The model in Fig. 2 was generated by a more careful docking of the acceptor stem with the productive subunit and suggests that the anticodon arm lies in the plane of the subunits, where it would be in close proximity to the crystallographic island of density mentioned above (Fig. 2).

H.B. thanks the Royal Society and EMBO for long-term fellowships.

Received 8 October 1985; accepted 17 January 1986.

1. Fersht, A. R. *et al.* *Angew. Chem.* **23**, 467–473 (1984).
2. Waye, M. M. Y., Winter, G., Wilkinson, A. J. & Fersht, A. R. *EMBO J.* **2**, 1827–1829 (1983).
3. Winter, G., Koch, G. L. E., Hartley, B. S. & Barker, D. G. *Eur. J. Biochem.* **132**, 383–387 (1983).
4. Brown, R. S. *et al.* *Nucleic Acids Res.* **5**, 23–26 (1978).
5. Winter, G., Fersht, A. R., Wilkinson, A. J., Zoller, M. & Smith, M. *Nature* **299**, 756–758 (1982).
6. Fersht, A. R., Mulvey, R. S. & Koch, G. L. E. *Biochemistry* **14**, 13–18 (1975).
7. Carter, P., Bedouelle, H. & Winter, G. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
8. Robertus, J. D. *et al.* *Nature* **250**, 546–551 (1974).
9. Bhat, T. N., Blow, D. M., Brick, P. & Nyborg, J. *J. molec. Biol.* **158**, 699–709 (1982).
10. Jones, T. A. in *Computational Crystallography* (ed. Sayre, D.) 303–310 (Clarendon, Oxford, 1982).
11. Wilkinson, A. J., Fersht, A. R., Blow, D. M. & Winter, G. *Biochemistry* **22**, 3581–3586 (1983).
12. Fersht, A. R. *Enzyme Structure and Mechanisms* (Freeman, New York, 1985).
13. Fersht, A. R. *Biochemistry* **14**, 5–12 (1975).
14. Carter, P., Bedouelle, H., Waye, M. M. Y. & Winter, G. *Oligonucleotide Site-directed Mutagenesis in M13: an Experimental Manual* (Anglian Biotechnology, Colchester, 1985).
15. Carter, P., Bedouelle, H. & Winter, G. *Nucleic Acids Res.* **13**, 4431–4443 (1985).
16. Wilkinson, A. J., Fersht, A. R., Blow, D. M., Carter, P. & Winter, G. *Nature* **307**, 187–188 (1984).
17. Barker, D. G. *Eur. J. Biochem.* **125**, 357–360 (1982).

Upstream activator sequences are present in the promoters of nitrogen fixation genes

Martin Buck*, Stephen Miller,
Martin Drummond & Ray Dixon

AFRC Unit of Nitrogen Fixation, University of Sussex,
Brighton BN1 9RQ, UK

Nitrogen fixation (*nif*) genes in *Klebsiella pneumoniae* are controlled by a transcriptional activation mechanism which requires the *ntfA* gene product acting concert with either the *ntfC* or *nifA* products (reviewed in ref. 1). The *nif* promoters lack the canonical -35, -10 promoter elements and instead their activation requires nucleotide sequences located around -24 and -12 (refs 2-4). We have now identified and characterized a further essential promoter sequence in the *K. pneumoniae nifH*, *nifU*, *nifB* and *ORF* promoters, located more than 100 base pairs (bp) from the transcription start site. This promoter element is required for *nifA*- but not *ntfC*-mediated activation and for the inhibition of chromosomal *nif* expression observed when cells harbour multiple copies of certain *nif* promoters⁵. The upstream sequence is conserved among 10 *Rhizobium* and 2 *Azotobacter nif* promoters. We show that the positioning and orientation of the upstream sequence is not critical for promoter activity up to a distance of 2 kilobases

(kb) and that the upstream sequence is itself transcriptionally inactive, probably acting in *cis* with the downstream sequences to produce a fully active promoter. Thus, an extended promoter structure is required for *nif* gene expression.

Deletion of nucleotides -72 to -184 in the *K. pneumoniae nifH* promoter greatly reduces its transcriptional activation by the *nif*-specific activator *nifA* (ref. 4), and relieves the inhibition of chromosomal *nif* gene expression² usually shown by multiple copies of the *nifH* promoter⁵. We have analysed in detail the role of upstream sequences in *nifH*, *nifU* and *nifB* promoter function. Figure 1a shows that upstream sequences are essential for the multicopy inhibitory effect of the *nifH* (pMB1, pSMM4) and *nifU* (pMD1106, pSMM9) promoters (the *nifB* promoter clone shows no multicopy inhibition). The *nifH* promoter can be weakly activated by *ntfC* (ref. 4), but we were unable to detect *ntfC* activation of the *nifU* and *nifB* promoters. The *nifH*, *nifU* and *nifB* promoters require upstream sequences for efficient *nifA* activation, but *ntfC* activation of the *nifH* promoter is independent of the upstream element. Therefore, the upstream element is required only for *nifA* activation, thereby indicating that the deletions do not disrupt a more general property of the promoter, for example secondary RNA polymerase binding sites⁶.

Within sequences deleted from the *nifH*, *nifU* and *nifB* promoters, we have identified a new *nif* consensus sequence that occurs upstream in *K. pneumoniae*, *Rhizobium* and *Azotobacter nif* promoters (Table 1). These sequences are delineated to within 8 bp at the 5' end and 4 bp at the 3' end by the deletion

Table 1 Upstream sequences in *nif* promoters

<i>nifH</i> Kp(16)	G C A G G C G C A A T T G T T C T G T T T C C C A C A T T T G
<i>nifU</i> Kp†	C G T G T C G T T T C T G T G A C A A A G C C C A C A A A A C
<i>nifB</i> Kp‡	A C T G T G C C T T A T G T G A G A T T C A G G A C A T T T G
<i>ORF</i> Kp(16)	T G G T G G C G C T T T G T C G C C T G T T C A A C A G C T T
<i>nifE</i> Kp†	C C A G C G C G A T A G G T C A T A A A G C A T C A C A T G C C
<i>nifH</i> Rm(17)	A C T G T A G C C C T T G T C G G C T T A G C G A C A C G A G
<i>nifH</i> Rt(18)	G T C G T C A C C T T T G T C G G C T T C G T G A C A C G C T*
<i>nifH</i> PR(19)	G G T G T G G C C G T T G T T C G T T T T C T G A C A G C C G
<i>nifDK</i> PR(20)	G C A A A C G C C G A T G T C G C C T T C G C A A C A A C A A
<i>nifH</i> Rp(21)	C G C T T T G T C T G A T C G G C G A C A T T A G*
<i>nifH</i> Rj(22)	T T G C G A T T G T T T G T T C G T T G T C T G A C A G C C G
<i>nifH</i> Rj(22)	C G G G C A G A T C T T G T C A G A T C C A A A A C A G C C T
<i>nifDK</i> Rj(23)	G C A A A G G C G A G T G T C G C A T T C G C A A C A A C A G
<i>nifDK</i> Cowpea R(24)	C C A A A G C G T G A T G T C G G G T T C G C A A C A A A T C
<i>IRC78</i>	T C C G T C G C C T C T G T C G G C C C T C G A C A C A A G
<i>Rm P₂</i> (25)	T T C G T C A C C T T T G T C G G C T T A A C C A C A C A A G
<i>Rm P₃</i> (25)	G A T G T A G G A C A T G T C G T G T A G C A T A C A C A G A
<i>nifH</i> Ac.§	C A T T G T T C G G T T G T A G C A A T T A C A C A C A T G T*
<i>nifH</i> Av.(26)	G G A C G C C T G C T T G T T G C A A A C C T G A C A G G A A*
<i>nifE</i> Av.(27)	
Consensus	G N ₇ T G T N ₄ T N ₅ A C A
Binding site consensus ⁷	T G T G T N ₆₋₁₀ A C A C A

For comparison, the position of the invariant G of the TGT motif is given, with the C residue of the invariant G-C dinucleotide promoter element designated as -12. Abbreviations used: Kp, *Klebsiella pneumoniae*; Rm, *Rhizobium meliloti*; Rt, *Rhizobium trifolii*; PR, *Parasponium rhizobium*; Rj, *Rhizobium japonicum*; Rp, *Rhizobium phaseoli*; Ac., *Azotobacter chroococcum*; Av., *Azotobacter vinelandii*. The upstream sequence is absent from the *K. pneumoniae nifL* promoter¹⁰. The *nifE* promoter shows partial homology only. A potential upstream promoter element exists in the *nifF* promoter¹⁰, but at position -263. The upstream element is present in two out of three *Rhizobium phaseoli nifH* promoter sequences²¹. The two potential upstream elements present in the *Rhizobium japonicum nifH* promoter sequence²² are shown. The upstream consensus sequence is also found in all three *Rhizobium trifolii* repeated sequence copies described in ref. 28. Both Ac. *nifH* and Av. *nifE* promoters have a TGT-N₁₀-ACA upstream sequence; the Av. *nifH* upstream sequence encodes a partially homologous element of the form TGT-N_{9/11}-ACA. References are shown in parentheses.

* Transcription start site has not been determined, but the promoter has been identified by homology with the *nif* consensus -24, -12 promoter sequences.

† J. Beynon, V. Buchanan-Wollaston, M. C. Cannon and F. C. Cannon, in preparation. ‡ M.D., unpublished data. § R. A. Jones, P. Woodley and R. C. Robson, unpublished data.

end points in plasmids pSMM8 and pSMM9 (Figs 3 and 1a, respectively). The conserved sequence conforms to a consensus sequence for protein binding sites on DNA⁷ and could be a site at which *nifA* interacts with *nif* promoters. The DNA sequence of the *nifA* gene predicts a potential DNA binding domain in the C-terminal region of the *nifA* protein⁸. The upstream promoter elements identified by us may interact with the *nifA* protein at this domain. The proposed binding sites lie more than 100 nucleotides from the transcription start (where this has been determined) and are characterized by an invariant TGT-N₁₀-ACA motif flanked by A+T-rich sequences.

Multiplicity inhibition is thought to result from activator titration⁹. We reasoned that, as the upstream consensus conforms to a protein binding site, and as a point mutation in the *nifH* promoter (G→T-136)^{2,4} within this sequence relieves multiplicity inhibition, *nifA* might bind to the upstream sequence. However, when we assessed the ability of the *nifH* upstream sequence alone to cause multiplicity inhibition (Fig. 1b), no inhibition was detected, even when the upstream element was cloned into the very high copy number pEMBL8⁺ plasmid. Furthermore, no inhibition was detected from subclones of the *nifH* promoter which retained only the downstream *nif* consensus promoter

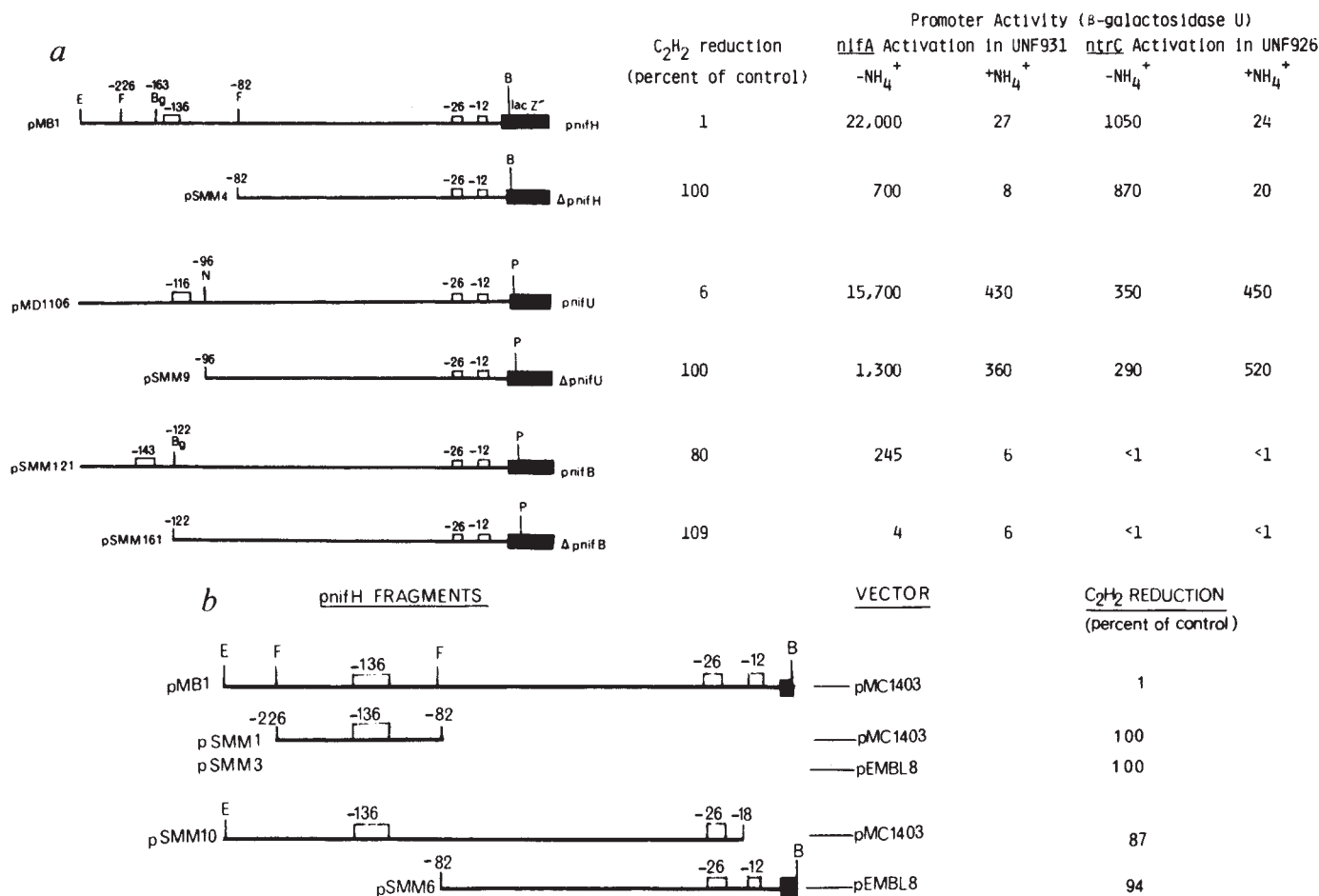


Fig. 1 **a**, Influence of upstream deletions on the activities of the *nifH*, *nifU* and *nifB* promoters. Each promoter clone (pnif) and the upstream deleted derivative (Δ pnif) was assayed for its multiplicity inhibitory effect on chromosomal *nif* expression by measuring the acetylene reduction activity of *K. pneumoniae* UNF931 (*ntr*⁺ *nif*⁺)⁵ carrying pnif and Δ pnif. These data are expressed as a percentage of the acetylene reduction activity of UNF931 carrying the promoter cloning vehicle, pMC1403. Transcriptional activation of pnif and Δ pnif by *nifA* was assayed in UNF931 and activation by *ntrC* in UNF926 (*ntr*⁺ Δ *his*-*nif*) as described previously⁴. The activators *nifA* and *ntrC* are functionally related proteins and can substitute for one another in the activation of a number of *ntr*-regulated promoters¹. Activation data are given as units of β -galactosidase. Open boxes at -136, -116 and -143 indicate the position of a conserved upstream sequence (see Table 1). The low level of activator-independent constitutive expression from the *nifU* promoter clones may be attributed to the sequence 5'-AATACA-N₁₆-ATTAAT behaving as a typical prokaryotic promoter of the -10, -35 form²⁹. Relevant restriction sites shown are: B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; F, *Fnu*DII; N, *Nru*I; P, *Pst*I. **b**, Multiplicity inhibition from subclones of the *nifH* promoter.

Methods. **a**, Plasmids: *nif*-*lacZ* translational fusions were constructed in the expression vector pMC1403 (ref. 30) or its derivative pMD1405 (ref. 4). The *nifH*-*lacZ* fusion plasmid pMB1 has been described previously⁴. Plasmid pSMM4 was constructed by cloning the downstream *Fnu*DII-*Bam*HI fragment from pMB1 into *Sma*I-*Bam*HI-digested pMC1403. Plasmids pMD1106 and pGS1103 are *nifU* and *nifB* translational fusions constructed in pMD1045. Deletion of the upstream sequence from the *nifU* promoter was achieved by isolating the shortened promoter as a *Nru*I-*Pst*I fragment and ligating this into *Sma*I-*Pst*I-digested pMD1106 (to retain the reading frame) to yield plasmid pSMM9. Plasmids pSMM121 and pSMM161, which carry the complete and upstream deleted *nifB* promoter respectively, were constructed by isolating the *nifB* promoter as an *Eco*RI-*Pst*I fragment from pGS1103, digesting this with *Eco*RV or *Bgl*I, ligating into *Sma*I-*Pst*I-digested pMD1405, treating with T4 DNA polymerase and religating. **b**, Assays were performed as for **a**. Plasmid pSMM1 was constructed by subcloning the *Fnu*DII restriction fragment carrying the upstream sequence of the *nifH* promoter into *Sma*I-digested pMC1403. The upstream sequence was also cloned into pEMBL8⁺ as an *Eco*RI-*Bam*HI fragment from pSMM1, as was the downstream sequence from pSMM4, to yield plasmids pSMM3 and pSMM6, respectively. Sequencing single-stranded template derived from pSMM3 showed that the upstream sequence in pSMM1 was in the same orientation as in pMB1. Plasmid pSMM10 was constructed by restricting the *nifH* *Eco*RI-*Bam*HI promoter fragment from pMB3 with *Rsa*I and ligating this material into *Eco*RI-*Sma*I-digested pMC1403. (Plasmid pMB3 carries a silent point mutation at -18 which creates an *Rsa*I restriction site⁴.)

Table 2 Activities of *nifL* and *nifL-nifH* hybrid promoters

Plasmid	Upstream sequence and orientation	Plasmid in <i>trans</i>			Acetylene reduction (%)
		None	pMC71A (<i>nifA</i> ^c) Activity	pMM14 (<i>ntnC</i> ^c) (β -galactosidase units)	
pRD531	<i>nifL</i>	50	2,200	2,900	—
pRD535	Δ <i>nifL</i>	70	430	340	100
pSMM535	<i>ORF</i> + <i>nifH</i>	10	2,600	26	50
pSMM536	<i>nifH</i> + <i>ORF</i>	11	1,300	130	37

Activation of *nifL* and *nifH-nifL* hybrid promoters by *nifA* and *ntnC* plasmids. Cultures were grown and assayed as described previously⁴. The strain background for activation assays was *Escherichia coli* ET8894 (Δ (*rha-ntnC*)1703 *rbs gyrA hutC lacZ*::IS1 Mucts62). Plasmids pMM14 and pMC71A provide *ntnC* and *nifA* in *trans* constitutively. Multicopy inhibition was assayed in *K. pneumoniae* UNF931 (ref. 4).

Table 3 Influence of spacing and orientation of upstream sequences on activation

Plasmid	Position of upstream elements	Element and orientation (turns of helix)	Activity (% pMB1)	C ₂ H ₂ reduction (% control)
pMB43	-102	$\overline{\text{H}} + \overline{\text{ORF}}$ (3.2)	44	75
pMB42	-103	$\overline{\text{ORF}} + \overline{\text{H}}$ (3.14)	105	48
pSMM48	-124	$\overline{\text{H}}$ (1.14)	50	8
pMB1 (wild type)	-136	$\overline{\text{ORF}} + \overline{\text{H}}$ (0)	100	1
pSMM8	-136	$\overline{\text{H}}$ (0)	79	0.3
pSMM15	-155	$\overline{\text{H}} + \overline{\text{ORF}}$ (1.81)	55	5
pSMM14	-156	$\overline{\text{ORF}} + \overline{\text{H}}$ (1.9)	46	1
pSMM47	-156	$\overline{\text{H}}$ (1.9)	56	3
pSMM142	-168	$\overline{\text{ORF}} + \overline{\text{H}}$ (3.04)	145	10
pSMM143	-190	$\overline{\text{ORF}} + \overline{\text{H}}$ (5.14)	51	70
pSMM144	-196	$\overline{\text{ORF}} + \overline{\text{H}}$ (5.71)	34	70
pSMM146	-200	$\overline{\text{ORF}} + \overline{\text{H}}$ (6.1)	27	68
pJMW31	-208	$\overline{\text{ORF}} + \overline{\text{H}}$ (6.9)	34	100
pSMM141	-1,350	$\overline{\text{ORF}} + \overline{\text{H}}$ —	12	100
pJMW3	-2,150	$\overline{\text{ORF}} + \overline{\text{H}}$ —	10	100
pSMM4	—	— —	3	100

See Fig. 3 legend.

sequence (Fig. 1b), indicating that the upstream element acts in *cis* with sequences at -24 and -12. Evidence for this comes from the finding that when cloned as an *Eco*RI fragment from pSMM13 into pACYC184, and present in *trans* to the upstream deleted plasmid pSMM4, there is no increase in the level of expression or multicopy inhibition from pSMM4 (data not shown). Transcriptional activity of the *nifH* upstream sequence was assessed by cloning it as an *Eco*RI-*Bam*HI fragment from pSMM1 into the transcriptional fusion vector pJEL126. Assays in UNF931 ($\pm$ NH₄) demonstrated that the *nifH* upstream element was transcriptionally inactive in the direction of *nifH* (data not shown). Therefore, all *nifH* transcription is initiated down-

stream and the upstream element cannot act as an independent promoter, but must act with the downstream consensus promoter sequences at -24 and -12 to facilitate multicopy inhibition and efficient transcription.

Upstream sequences are necessary for full expression of the *nifL* promoter but, unlike the *nifH* promoter, these sequences show no specificity towards *ntnC* versus *nifA* activation¹⁰. Consistent with this is the finding that the *nifL* promoter lacks sequences shown in Table 1. To examine the *nifA*-specific character of the upstream element, a hybrid promoter was constructed carrying the *nifL* downstream elements and upstream sequences from the *nifH* promoter (Fig. 2). Other constructs (see below) had demonstrated that the spacing was not critical for activation. Results (Table 2) show that the *nifH* upstream sequence restored efficient *nifA* activation to a *nifL-lac* promoter (pRD535) lacking *nifL* upstream sequences¹⁰. Although *nifA* activation was enhanced, *ntnC* activation remained low, confirming the *nifA*-specific character of the upstream element. The hybrid promoters (pSMM535 and 536) inhibited acetylene reduction in UNF931, consistent with the observation that the upstream element is required in the *nifH* and *nifU* promoters for multicopy inhibition and acts in *cis* (Fig. 1).

We constructed several mutant derivatives of the *nifH* promoter in which the spacing between the upstream and downstream elements was altered (Fig. 3, Table 3). The wild-type *nifH* promoter carries an upstream element believed to be associated with the *ORF* promoter (Fig. 3, Table 1). Starting with plasmid pSMM8, which lacks the *ORF* upstream element, we examined both the influence of the *ORF* upstream element and the orientation of the *nifH* upstream element on expression (Fig. 3), and found (Table 3) that: (1) the *ORF* sequence is not required for *nifH* promoter expression (plasmids pSMM8,

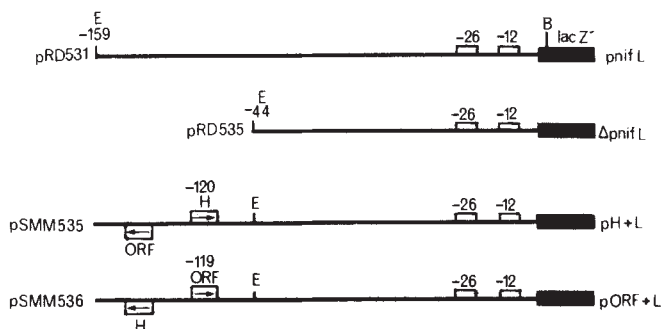


Fig. 2 Construction of hybrid *nifH-nifL* promoters. The upstream element from the *nifH* promoter clone pMB1, which carries both *nifH* and *ORF* upstream sequences, was cloned from pSMM13 (see Fig. 3 legend) as an *Eco*RI restriction fragment into *Eco*RI-linearized pRD535. Plasmid pSMM535 carries the *nifH* upstream sequence proximal to *pnifL*, pSMM536 carries the *ORF* upstream sequence proximal to *pnifL*. The position of the G residue of the TGT motif is given for plasmids pSMM535 and pSMM536.

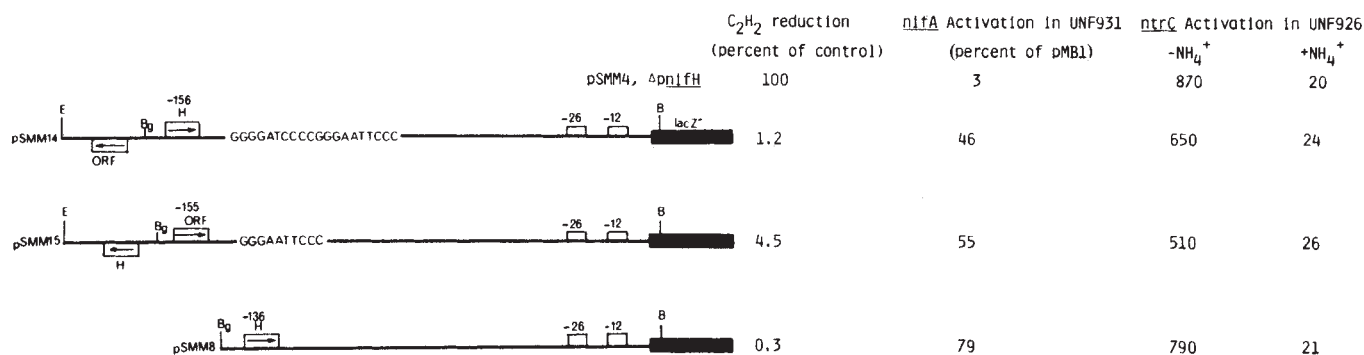


Fig. 3 Influence of spacer mutations and the *ORF* upstream sequence on *nifH* promoter activity (see also Table 3). Assays of activity were performed as for Fig. 1a, with *nifA* activation data being presented as a percentage of the pMB1 activity. Data from the upstream *nifH* promoter deletion plasmid pSMM4 ($\Delta nifH$) are included for comparison. The position of the G residue of the TGT motif is indicated, and the number of turns of the helix by which this is displaced from the wild-type position at -136 is given. Plasmids pSMM14 and pSMM15 were constructed by first cloning the upstream sequence from pSMM1 into pEMBL8⁺ on a *Pst*I-*Bam*HI fragment to yield pSMM13, flanking the upstream element as an *Eco*RI restriction fragment and ligating this in both orientations into pSMM4 linearized using *Eco*RI. The additional DNA gained from these cloning steps is shown. The *ORF* upstream sequence was deleted from pMB1 by treating the *Bgl*II fragment carrying the *nifH* promoter (and part of *lacZ*) with T4 DNA polymerase, cutting with *Bam*HI and ligating into *Sma*I-*Bam*HI-digested pMC1403 to yield pSMM8. To obtain the *nifH* upstream sequence on an *Eco*RI fragment in the absence of the *ORF* upstream element, the *Eco*RI-*Fnu*DII fragment carrying only the *nifH* upstream sequence was subcloned from pSMM8 into *Eco*RI-*Sma*I-digested pMC1403 to yield pSMM81. The *nifH* upstream sequence was then subcloned from pSMM81 as a *Pst*I-*Bam*HI fragment into pEMBL8⁺ to yield pSMM82, permitting its isolation as an *Eco*RI fragment which was then ligated in both orientations into *Eco*RI-linearized pSMM4 to yield plasmids pSMM47 and pSMM48. To increase the spacing between upstream and downstream elements, the kanamycin resistance gene from pACYC177, flanked by an inverted repeat of cloning nests from pEMBL18 (M.D., unpublished), was cloned as a *Hinc*II fragment into the *Sma*I site of pSMM14 to yield pSMM41. Subsequent deletion of the kanamycin resistance gene by restriction with *Xba*-1, *Sma*-1 or *Asp*-718 yielded plasmids pSMM142, pSMM143 and pSMM144, respectively. Filling in the *Asp*-718 site in pSMM144 using Klenow polymerase yielded pSMM146. Cloning the Ω insertion element³¹ as a *Sma*I fragment from pHB34 Ω into the *Sma*I site of pSMM14 yielded pJMW3, and subsequent deletion of the insertion element using the *Hind*III sites yielded pJMW31. Plasmids pMB42 and pMB43 were obtained by cloning the *Eco*RI fragment carrying the upstream sequences from pSMM13 into pMB41. Plasmid pMB41 is deleted for nucleotides through to position -27 and carries a silent T to C transition at -27, being derived from pMB4 (ref. 4). For all upstream spacer mutants, equivalent derivatives of pSMM4 were constructed, thus carrying the packing sequences but no upstream elements. In all cases, derepression assays of these control plasmids in UNF931 produced 400–600 units of β -galactosidase activity.

pSMM47 and pSMM48), (2) the upstream element can activate transcription even when placed 2 kb away from the downstream promoter elements, but has an optimal position around position -136 and (3) when the orientation of the *nifH* upstream element is reversed, it is still active (plasmids pSMM8, pSMM47 and pSMM48). In the last case, the sequence of the element is 5'-CCAAATGTGGGAAACAGAACAAATTGCGC, retaining only those underlined sequences in common with the element in its usual orientation (Table 1). Nucleotides -28 to -80 of the *nifH* promoter are absent from plasmids pMB42 and pMB43, demonstrating that most nucleotides between the upstream and downstream elements are not essential for activation. Multicopy inhibition assays (Table 3) indicate that this parameter of *nifH* promoter activity is more sensitive to changes in the spatial relationship between upstream and downstream elements than is activation. However, the variability associated with the multicopy inhibition assay prevents us from determining whether a low level of inhibition can be attributed to plasmids pSMM141, pJMW3 and pJMW31.

We have suggested that the upstream promoter element is a site at which *nifA* interacts to promote transcription. For several positively controlled promoters, the proximity of activator and RNA polymerase binding sites suggests that transcriptional activators can interact directly with RNA polymerase¹¹. Because of the greater distance over which the upstream element can act (Fig. 3), such an interaction seems unlikely in the case of *nif* promoters. Rather, we suggest that if a *nifA*-polymerase interaction is necessary for activation, it could occur either through a tertiary interaction facilitated by a folding in the DNA structure (rather than as a linear protein arrangement), as has been proposed for *araC*- and *galR*-mediated repression^{12,13}, or through *nifA* interacting with the upstream sequence and then sliding along the DNA towards proteins located at the downstream sequences. The folding model may require that the faces of the helix containing upstream and downstream sequences

bear a strict spatial relationship to each other in order that proteins interacting with these sequences make contact correctly. However, it is not possible at present to distinguish rigorously between the various mechanistic possibilities.

The upstream sequence seems to reflect a specialization among certain *nif* promoters, not restricted to *K. pneumoniae*, permitting efficient activation by regulatory proteins functionally homologous to *nifA*. Recently, Better *et al.*¹⁴ observed that upstream sequences were required for maximal activation of the *Rhizobium meliloti nifHDK* promoter by the *K. pneumoniae nifA* gene product. The promoter-down phenotype of these upstream deletions¹⁴ could be attributed to the loss of the *R. meliloti* upstream promoter element (Table 1). A similar requirement for upstream sequences has been found in the expression of the *Rhizobium japonicum nifH* and *nifD* promoters (A. Alvarez-Morales and H. Hennecke, personal communication). It is interesting to note the similarities between the properties of the upstream DNA sequence required for activation and those of eukaryotic enhancer sequences, in particular the ability to function over a considerable distance and the independence of orientation¹⁵.

We thank J. Beynon and F. Cannon for making available unpublished *K. pneumoniae nif* sequence data, R. Robson for the *A. chroococcum nifH* sequence, W. Cannon for technical support, J. Woodcock for constructing plasmids pJMW3 and pJMW31, M. Merrick for useful discussions, and J. R. Postgate for his critical reading of the manuscript.

Received 16 December 1985; accepted 5 February 1986.

1. Dixon, R. A. *J. gen. Microbiol.* **130**, 2745–2755 (1984).
2. Brown, S. E. & Ausubel, F. M. *J. Bact.* **157**, 143–147 (1984).
3. Ow, D. W., Xiong, Y., Gu, Q. & Shen, S.-C. *J. Bact.* **161**, 868–874 (1985).
4. Buck, M., Khan, H. & Dixon, R. *Nucleic Acids Res.* **13**, 7621–7638 (1985).
5. Buchanan-Wollaston, V., Cannon, M. C. & Cannon, F. C. *Molec. gen. Genet.* **184**, 102–106 (1981).
6. Travers, A. A., Lamond, A. I., Mace, H. A. F. & Berman, M. L. *Cell* **35**, 265–273 (1983).

7. Gicquel-Sanzey, B. & Cossart, P. *EMBO J.* **1**, 591-595 (1983).
8. Drummond, M., Whitty, P. & Wootton, J. *EMBO J.* **5**, 441-447 (1986).
9. Riedel, G. E., Brown, S. E. & Ausubel, F. M. *J. Bact.* **153**, 45-56 (1983).
10. Drummond, M., Clements, J., Merrick, M. & Dixon, R. *Nature* **301**, 302-307 (1983).
11. Raibaud, O. & Schwartz, M. A. *Rev. Genet.* **18**, 173-206 (1984).
12. Dunn, T. M., Hahn, S., Ogden, S. & Schlieff, R. F. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5017-5020 (1984).
13. Majumdar, A. & Adhya, S. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6100-6104 (1984).
14. Better, M., Ditta, G. & Helinski, D. *EMBO J.* **4**, 2419-2424 (1985).
15. Dynan, W. S. & Tjian, R. *Nature* **316**, 774-778 (1985).
16. Scott, K. F., Rolfe, B. G. & Shine, J. *J. molec. appl. Genet.* **1**, 71-81 (1981).
17. Sundaresan, V., Jones, J. D. G., Ow, D. W. & Ausubel, F. M. *Nature* **301**, 727-732 (1983).
18. Scott, K. F., Rolfe, B. G. & Shine, J. *DNA* **2**, 149-155 (1983).
19. Scott, K. F., Rolfe, B. G. & Shine, J. *DNA* **2**, 141-148 (1983).
20. Weinman, J. J., Fellows, F. F., Gresshoff, P. M., Shine, J. & Scott, K. F. *Nucleic Acids Res.* **12**, 8329-8344 (1984).
21. Quinto, C. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 1170-1174 (1985).
22. Fuhrmann, M. & Hennecke, H. *J. Bact.* **158**, 1005-1011 (1984).
23. Kaluza, K. & Hennecke, H. *Molec. gen. Genet.* **196**, 35-42 (1984).
24. Yum, A. C. & Szalay, A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7358-7362 (1984).
25. Better, M., Lewis, B., Corbin, D., Ditta, G. & Helinski, D. R. *Cell* **35**, 479-485 (1983).
26. Brigle, K. E., Newton, W. E. & Dean, D. R. *Gene* **37**, 37-44 (1985).
27. Dean, D. R. & Brigle, K. E. *Proc. natn. Acad. Sci. U.S.A.* **82**, 5720-5723 (1985).
28. Schofield, P. R. & Watson, J. M. *Nucleic Acids Res.* **13**, 3407-3418 (1985).
29. Hawley, D. K. & McClure, W. R. *Nucleic Acids Res.* **11**, 2237-2255 (1983).
30. Casadaban, M. J., Chou, J. & Cohen, S. N. *J. Bact.* **143**, 971-980 (1980).
31. Prentki, P. & Krisch, H. M. *Gene* **29**, 303-313 (1984).

Tertiary structural similarity between a class A β -lactamase and a penicillin-sensitive D-alanyl carboxypeptidase-transpeptidase

B. Samraoui*†‡, B. J. Sutton*†, R. J. Todd*†,
P. J. Artymink*‡, S. G. Waley§ & D. C. Phillips*||

* Laboratory of Molecular Biophysics, Rex Richards Building,
South Parks Road, Oxford OX1 3QU, UK

† Laboratoire pour l'utilisation du Rayonnement Electromagnétique
(LURE), CNRS-Université Paris Sud, 91405 Orsay, France

§ Sir William Dunn School of Pathology, South Parks Road,
Oxford OX1 3QU, UK

β -Lactam antibiotics—the penicillins, cephalosporins and related compounds—act by inhibiting enzymes that catalyse the final stages of the synthesis of bacterial cell walls^{1,2}. Recent crystallographic studies of representative enzymes^{3,4} are beginning to reveal the structural bases of antibiotic specificity and mechanism of action, while intensive efforts are being made to understand the β -lactamase enzymes that are largely responsible for bacterial resistance to these antibiotics⁵⁻⁸. It has been suggested that the β -lactamases and β -lactam target enzymes may be evolutionarily related⁹ and some similarity of amino-acid sequence around a common active-site serine residue supports this idea¹⁰. We present here the first evidence from a comparison of three-dimensional structures in support of this hypothesis: the structure of β -lactamase I from *Bacillus cereus*¹¹ is similar to that of the penicillin-sensitive D-alanyl-D-alanine carboxypeptidase-transpeptidase from *Streptomyces* R61¹³.

β -Lactamase I (EC 3.5.2.6) from *B. cereus* 569 is a member of class A, according to the classification of β -lactamases on the basis of amino-acid sequence^{12,13}. It was isolated and purified following Baldwin *et al.*¹⁴ and crystals were grown by a vapour-diffusion technique; although small, they diffract to high resolution¹⁵. They have unit-cell dimensions $a = 143.0$, $b = 35.8$, $c = 57.2$ Å; $\beta = 97.8^\circ$. The space group is C2, with one molecule of β -lactamase I per asymmetric unit and about 50% by volume liquid of crystallization. Electron-density maps of the crystal structure have been obtained at 6 Å and at higher resolutions

up to 2.5 Å by the method of multiple isomorphous replacement combined with the use of anomalous scattering (MIRAS). Three useful heavy-atom derivatives were obtained by soaking the crystals in solutions of $K_2Pt(C_2O_4)_2$, $Sm(CH_3COO)_3$ and CH_3HgCl . X-ray diffraction data were collected photographically and, because of the very small size of the crystals, mainly by the use of synchrotron radiation at LURE, France, and at the Synchrotron Radiation Laboratory of the Science and Engineering Research Council, Daresbury. The position of the heavy atom in the Pt derivative was readily derived from isomorphous and anomalous difference Patterson maps and the heavy-atom positions in the other derivatives were obtained from difference-Fourier maps based on Pt phases. The FHLE [F_H (lower estimate)] and phase-refinement methods¹⁶ were used to refine the heavy-atom parameters, giving the values at 2.5 Å resolution shown in Table 1. The mean figure of merit at 6 Å resolution was 0.86 and that at 2.5 Å resolution, 0.66.

At low resolution, the boundary of the protein molecule is clear and the protein density includes the rod-like features characteristic of α -helices (Fig. 1). At 2.5 Å resolution, however, the electron density is disappointing, with relatively poor contrast. This is because the derivatives are not ideally isomorphous and is reminiscent of the difficulties encountered in studies of other β -lactamases^{5,6}. Taken together, our results prompt consideration of the suggestion that β -lactamases have variable conformations¹⁷. However, the map at high resolution was much improved by use of the Wang method of density modification¹⁸, applied to the MIR phases, which clarified much of the helical density and density corresponding to a β -pleated sheet, and this modified map provided a satisfactory basis for determination of the structure through the iterative process of map interpretation, restrained-least-squares refinement¹⁹, phase calculation and map recalculation based on combined MIRAS and calculated phases^{20,21}. The initial model, after least-squares refinement, gave a reliability index of 0.37 and this was reduced to 0.31 in one cycle of the iterative procedure. Figure 2 shows the five-stranded β -pleated sheet in the resultant combined-phase map. At this stage, the elements of secondary structure in the molecule are clearly defined, although the connections between them have not all been firmly established.

The overall dimensions of the molecule, excluding some external loops connecting secondary structures, are approximately $33 \times 38 \times 49$ Å. Figure 1a shows the arrangement of α -helices and β -pleated sheet in the molecule; it may be described as a five-stranded β -pleated sheet with a group of three α -helices on one side of it and five helices on the other. At this stage the molecule does not appear to consist of separate domains of the distinctness observed, for example, in phosphoglycerate kinase²², but regions of different structural types, both α and α/β , are present.

This arrangement of secondary-structure elements in β -lactamase I shows a striking resemblance to the structure of the penicillin-sensitive D-Ala-D-Ala carboxypeptidase-transpeptidase from *Streptomyces* R61¹³ (hereafter referred to as R61-CPase). Figure 1b shows the positions of helices and β -pleated sheet in this enzyme, derived from Kelly *et al.*²³ by the method of Rossmann and Argos²⁴, superimposed on the corresponding region of the low-resolution map of β -lactamase I. The arrangement of helices and sheet is clearly similar to that observed in the part of β -lactamase I shown as a heavy line in Fig. 1a and repeated, superimposed on the low-resolution electron density, in Fig. 1c. Helices A, B, C, D and H of the R61-CPase correspond well with helices in β -lactamase I. Furthermore, the general position of the β -pleated sheet is the same in the two molecules, although the sheet seems to be more extensive in β -lactamase I and there is a difference in the orientation of the strands of about 25° relative to the common helices. Three α -helical regions of the β -lactamase I do not have helical counterparts in the R61-CPase structure, and these lie above the section of electron density shown in Fig. 1c and are shown as a faint line in Fig. 1a.

† Present addresses: Department of Biochemistry and Molecular Biology, Harvard University, Massachusetts 02138, USA (B.S.); Department of Biochemistry, Sheffield University, Sheffield S10 2TN, UK (P.J.A.).

|| To whom correspondence should be addressed.

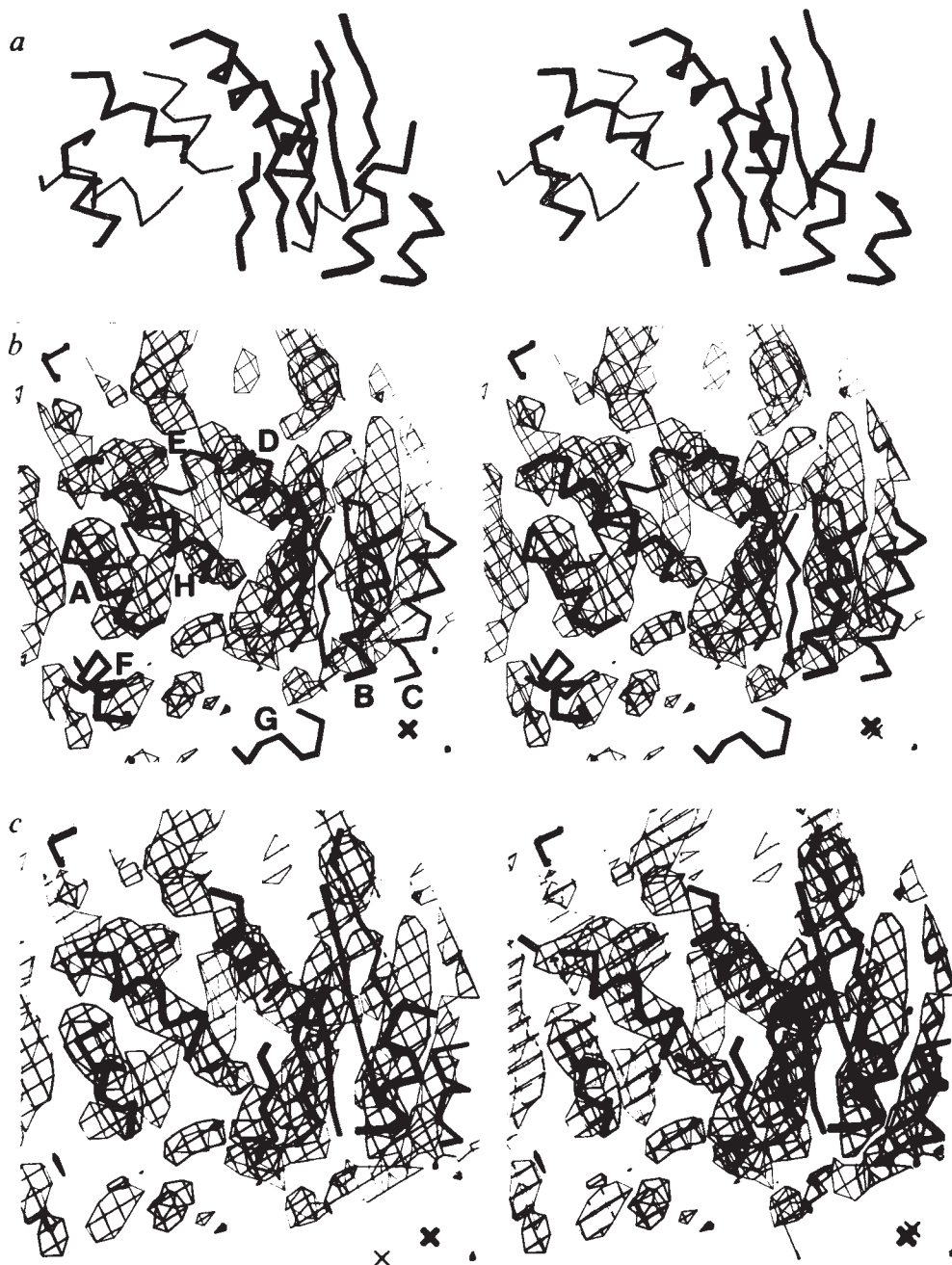


Fig. 1 *a*, The secondary-structure elements of β -lactamase I derived from analysis at 2.5 Å resolution. Those shown in faint line do not correspond to features in R61-CPase. *b*, A stereo view of part of the electron density in the 6-Å MIRAS map of β -lactamase I with secondary structure elements of the D-Ala-D-Ala carboxypeptidase-transpeptidase from *Streptomyces* R61²⁰ superimposed. The labelling of the helices is taken from ref. 3; in this more recent model of the structure, the position of helix G is slightly altered (together with the connections between the secondary structures, which are not considered here), but it is clear that there is no electron density in the map of β -lactamase I which corresponds to helices E, F and G of the R61-CPase. *c*, The electron density of *b* with an α -carbon atom trace of the structure of β -lactamase I derived at high resolution superimposed. Five of the eight α -helical segments can be seen in this section of the map and model, together with the five-stranded β -pleated sheet.

More detailed comparisons must await a complete interpretation of the β -lactamase map and refinement of both structures, but it is already apparent that there is no density in β -lactamase I corresponding to helices E, F and G in the R61-CPase. This is in accord with the prediction of Kelly *et al.*³, made on the assumption of sequence homology between the two molecules in the N-terminal region and the difference between their relative molecular masses, that β -lactamases of class A would lack a part of the carboxy-terminal domain of the R61-CPase. Interestingly, however, helix H, the carboxy-terminal helix of the R61-CPase, has a counterpart in β -lactamase I.

This discovery that a β -lactamase enzyme of class A and a penicillin-sensitive D-Ala-D-Ala carboxypeptidase-transpeptidase enzyme share an extensive region of common tertiary structure suggests very strongly that these two groups of enzymes have evolved by divergence from a common ancestor²⁵. This idea was suggested earlier by the weak but significant similarity between sequences surrounding the active-site serine residues

of the class A β -lactamases and two D-alanyl carboxypeptidases from *Bacillus* species¹⁰, and also the amino-terminal regions of penicillin-binding proteins from *Escherichia coli*²⁶. Apart from the fact that the R61-CPase is also a serine enzyme, however, the incomplete sequence data available appear to show no resemblance with the class A β -lactamases¹. The tertiary structural relationship reported here may therefore provide an example of the persistence of similarities between three-dimensional structures when the similarity between primary structures has disappeared²⁵.

R61-CPase shares some characteristics with the class C β -lactamases, which again are serine enzymes²⁷: for example, with respect to the relative rates of acylation and deacylation at the active serine residues. Whereas for β -lactamase I, both acylation by benzylpenicillin and deacylation are rapid, for R61-CPase acylation is more than a million times faster than deacylation² and the acyl enzyme accumulates, as it does for a class C β -lactamase²⁸. The relative molecular mass of R61-CPase is

Table 1 Refined heavy-atom parameters of the three derivatives for the final 2.5 Å-resolution phase calculation

	Scale factor	Relative occupancy	X	Y	Z	B	R_K
Ptb	1.0249	1.8608	0.1800	0.0000	0.2915	10.67	0.07
Ptc	1.0249	1.8559	0.1807	-0.0052	0.2914	3.78	0.08
Sm104	1.0319	0.9832	0.4977	0.3551	-0.0312	62.36	0.16
Sm105	1.0232	0.7653	0.4984	0.3570	0.0189	1.0†	0.19
Sm108	1.0178	1.3591	0.4992	0.3637	0.0194	122.29	0.12
Hg1	1.0270	1.3024	0.1777	0.0122	0.2882	21.77	0.13
Hg2	—	1.2987	0.1662	0.1443	0.4977	82.46	—
Hg3	—	0.4590	0.2315	0.1689	0.7361	55.47	—
Hg4	—	0.2778	0.2468	0.1879	0.8604	6.67	—

Overall mean figure of merit, 0.66.

$$R_K = \sum |F_{PH(obs.)} - F_{PH(calc.)}| / \sum F_{PH(obs.)}$$

where $F_{PH(obs.)}$ and $F_{PH(calc.)}$ are, respectively, the observed and calculated heavy-atom derivative structure factor amplitudes¹⁶. Ptb, Ptc, single-site derivative; two crystals, b^* and a^* axis data. Sm104, 105, 108, single-site derivative; three crystals, b^* axis data. Hg1–4, four-site derivative; one crystal, b^* axis data.

† Temperature factor fixed.

38,000 and in this respect also it is more similar to a class C (39,600) than to class A (29,000) β -lactamase. Furthermore, there is some similarity between the sequences immediately adjacent to the active-site serines of R61-CPase and class C β -lactamases³, as well as between a region beyond the active serine of a class C β -lactamase and the putative corresponding regions in PBP5 of *E. coli* and the membrane-bound CPase of *Bacillus subtilis*². There is reason to suspect, therefore, that the CPases and class C β -lactamases have a common evolutionary origin, although additional structural data are clearly needed to establish the relationship². However, no sequence similarity has been detected between the class A and class C β -lactamases and it has been concluded, therefore, that they evolved independently³. This conclusion must now be questioned. If the CPases are indeed homologous with the class C β -lactamases and if, as appears from the result reported here, they are also homologous with the class A β -lactamases, then the CPases, the class A and the class C β -lactamases must all have a common evolutionary origin. Consequently, the class C β -lactamases may be expected to share some or all of the tertiary structural features that we have found to be common to the class A β -lactamase I and the R61-CPase enzymes.

Finally, we note that we may shortly be able to compare the zinc-dependent β -lactamase II from *B. cereus*²⁹, the structure of which is being determined in this laboratory, with the zinc-dependent CPase from *Streptomyces albus* G (ref. 4). There is no sequence similarity between the two enzymes and the latter is only weakly sensitive to β -lactams, but the relationship suggested here between the serine-dependent β -lactamases and penicillin-sensitive enzymes prompts the thought that the zinc-dependent enzymes may also exhibit a family relationship.

We thank Sheila Gover, David Stuart, John Helliwell, Victor Inzani, V. Knott-Hunziker, Graham Knight and Joan Monks for their help with this work, Dr O. Dideberg for providing the Pt heavy-atom compound, and especially Dr R. Fourme, LURE and the SERC Daresbury Laboratory for their generous provision of synchrotron radiation facilities. This research was supported by the MRC, the E.P. Abraham Cephalosporin Trust (R.J.T.), the Algerian Government (B.S.), the Royal Society (B.J.S.) and the Science and Engineering Research Council. This is a contribution from the Oxford Enzyme Group.

Received 2 December 1985; accepted 7 February 1986.

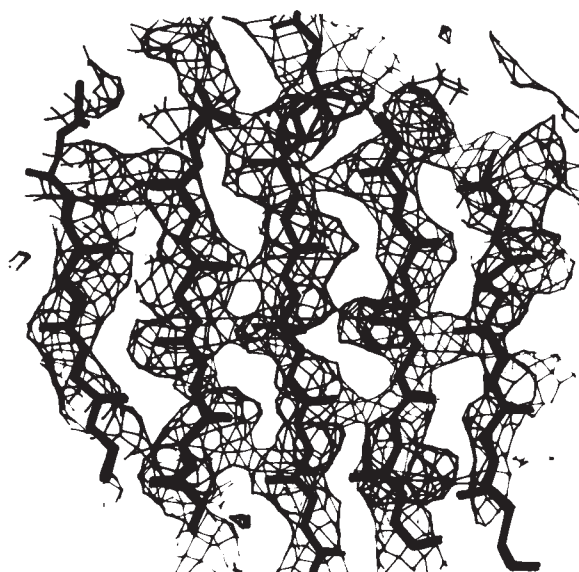


Fig. 2 Part of the electron-density map of β -lactamase I at 2.5 Å resolution after preliminary refinement. Five strands of β -pleated sheet are shown.

- Frère, J.-M. & Joris, B. *CRC crit. Rev. Microbiol.* **11**, 299–396 (1985).
- Waxman, D. L. & Strominger, J. L. *A. Rev. Biochem.* **52**, 825–869 (1983).
- Kelly, J. A. *et al. J. biol. Chem.* **260**, 6449–6458 (1985).
- Dideberg, O. *et al. Nature* **299**, 469–470 (1982).
- Knox, J. R., Kelly, J. A., Moews, P. C. & DeLucia, M. L. in *Beta-Lactamases*, 127, (Academic, London, 1979).
- Moult, J. *et al. Biochem. J.* **225**, 167–176 (1985).
- Dideberg, O. *et al. J. molec. Biol.* **181**, 145–146 (1985).
- Charlier, P., Dideberg, O., Frère, J.-M., Moews, P. C. & Knox, J. R. *J. molec. Biol.* **171**, 237–238 (1983).
- Pollock, M. R. *Proc. R. Soc. B179*, 385–401 (1971).
- Waxman, D. J. & Strominger, J. L. *J. biol. Chem.* **255**, 3964–3976 (1980).
- Kuwabara, S. & Abraham, E. P. *Biochem. J.* **103**, 27c–30c (1967).
- Ambler, R. P. *Phil. Trans. R. Soc. B289*, 321–331 (1980).
- Jaurin, B. & Grundstrom, T. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4897–4901 (1981).
- Baldwin, G. S. *et al. Biochem. J.* **191**, 111–116 (1980).
- Aschaffenburg, R. *et al. J. molec. biol.* **120**, 447–449 (1979).
- Blundell, T. L. & Johnson, L. N. *Protein Crystallography* (Academic, London, 1976).
- Virden, R., Bristow, A. F. & Pain, R. H. *Biochem. biophys. Res. Commun.* **82**, 951–956 (1978).
- Wang, B. C. in *Diffraction Methods for Biological Macromolecules* (eds Wyckoff, H. W., Hirs, C. H. W. & Timasheff, S. N.) 90–112 (Academic, New York, 1985).
- Hendrickson, W. A. & Konert, J. H. in *Computing in Crystallography* (eds Diamond, R., Ramaseshan, S. & Venkatesan, K.) 13.1–13.26 (Indian Academy of Sciences, Bangalore, 1980).
- Stuart, D. I. & Artymuk, P. J. *Acta crystallogr.* **A40**, 713–716 (1984).
- Huber, R. *et al. J. molec. Biol.* **89**, 73–101 (1984).
- Banks, R. D. *et al. Nature* **279**, 773–777 (1979).
- Kelly, J. A., Moews, P. C., Knox, J. R., Frère, J.-M. & Ghuyssen, J.-M. *Science* **218**, 479–481 (1982).
- Rossmann, M. G. & Argos, P. *Acta crystallogr.* **B36**, 819–823 (1980).
- Phillips, D. C., Sternberg, M. J. E. & Sutton, B. J. in *Evolution from Molecules to Men* (ed. Bendall, D. S.) 145–173 (University of Cambridge Press, 1983).
- Waxman, D. J., Amanuma, H. & Strominger, J. L. *FEBS Lett.* **139**, 159–163 (1982).
- Coulson, A. *Biotechnol. gene. Engng Rev.* **3**, 219–253 (1985).
- Knott-Hunziker, V., Petrusson, S., Waley, S. G., Jaurin, B. & Grundstrom, T. *Biochem. J.* **207**, 315–322 (1982).
- Sabath, L. D. & Abraham, E. P. *Biochem. J.* **98**, 11c–13c (1966).

Estimation of scrapie nucleic acid MW from standard curves for virus sensitivity to ionizing radiation

IN an unusually rancorous communication to this forum¹, Alper protests Fig. 4 of my letter² in which I constructed standard curves for the inactivation of single-stranded (ss) or double-stranded (ds) viruses by ionizing radiation versus their nucleic acid molecular weights (MWs). Extrapolation of either line to the inactivation rate constant (D_{37}) of the scrapie virus gave virus-like values for the scrapie genome. This is in sharp contrast to values of 150,000 daltons or less calculated from the 'target theory' by Alper³ and others².

In defending her target calculation¹, Alper denounces my use of previous reviews (as a source of data for constructing these curves, incorrectly), asserting that "some of the references are wrong" and that few of the references were checked. She gives the following four examples:

Yellow fever virus (YFV). Alper: "Molecular weight too low, about one-tenth of what we took from our independent source. Rohwer's source makes no mention of the MW of this virus." The YFV RNA has been sequenced and has a MW of 3.7×10^6 daltons⁴, in good agreement with the value plotted in Fig. 4, 3.5×10^6 daltons (as given by) my source⁵ (on page 18). The "independent source"⁶ preferred by Alper is an ultrastructural study of intact YFV virions. Her value was apparently estimated from the virion core diameter.

Vaccinia. Alper: " D_{37} incorrectly quoted by Rohwer from K & M [Kaplan and Moses] who cited McCrea incorrectly." Kaplan and Moses⁷ give a range of values for vaccinia obtained from multiple determinations by McCrea⁸ and Wilson⁹, both of whom they cite. I plotted the mean of the range. Alper ignores Wilson⁹, selects the single value from McCrea⁸ most favourable to her argument, which is also the only determination that appears to have been made under vacuum, and then objects to the use of vacuum.

Shope papilloma. Alper: " D_{37} plotted as too high to fit the Lea theory. K & M's source is Syverton *et al.*, who gave the dose for 'total inactivation'—an unknown multiple of the inactivation dose...". Indeed Syverton *et al.*¹⁰ summarize their data in terms of total inactivation rather than inactivation rate constant. However, the rate constant is easily calculated from the data in experiments II-IV and this is the value cited by K & M and plotted in Fig. 4. The identical value was obtained by Lea *et al.*¹¹ in their analysis of the

completely independent measurements of Friedewald and Anderson¹².

Newcastle¹² disease virus (NDV). Alper: "Inactivation dose too low... Rubin and Temin... irradiated the virus in suspension." NDV has also been irradiated dry¹³, giving a rate constant even smaller than that plotted in Fig. 4. Most investigators, aware that viruses tolerate drying poorly, suspend viruses in broth, serum or tissue homogenates, either wet or frozen, to protect against indirect effects, and for this reason most of the available data have been obtained in suspension. While Alper insists that viruses be irradiated dry, she invokes this argument to discard only the particularly inconvenient NDV point from Fig. 4. If we apply her criteria consistently and consider only the five viruses irradiated dry in Fig. 4 (three from her own laboratory), both the line for the ss viruses and that for the ds viruses closely parallel the original regression lines in Fig. 4 predicting ds scrapie genome of $MW 2.0 \times 10^6$ daltons or a ss genome of 0.86×10^6 daltons. These results are consistent with a larger study to be published separately (manuscript in preparation), in which each state, liquid, frozen and dry, was considered separately.

Using these four highly contrived 'examples', Alper alleges that the integrity of Fig. 4 is compromised by the use of previous reviews in documenting the data. *Nature's* succinct format necessitated the use of reviews where possible and did not permit elaboration on individual data points in the two paragraphs of text devoted to Fig. 4.

Alper intimates that low-temperature inactivations may be in error. While it is true that some viruses¹⁴, but not others¹⁵, show enhanced protection at liquid nitrogen temperatures (-196°C), the only frozen specimens in Fig. 4 were at dry ice temperatures (-78°C), where the available evidence^{15,16}, including that for scrapie^{17,18}, suggests virus inactivation kinetics similar to those obtained dry.

Alper's most surprising suggestion is that ss and ds viruses (and even proteins) be regressed as a single group. This is in spite of overwhelming evidence that ss viruses show a 20–40-fold greater sensitivity to inactivation than do ds viruses of similar genome size^{7,19–24}. This difference derives in part from the ease with which many otherwise lethal single-strand lesions can be repaired in ds DNA.

As noted by Alper, crude tissue suspensions of the scrapie agent show greater sensitivities to inactivation at ultraviolet wavelengths below the 254 nm maximum typical of some viruses²⁵. However, a similar inactivation spectrum is observed for tobacco mosaic virus (TMV)²⁶ and, to a lesser extent, potato X virus²⁷. The absence of other examples may only reflect the limited number of viruses for which an inactivation spectrum has been determined rather than the actual prevalence of this presumed anomaly in nature. It may be noteworthy that TMV and potato X share a filamentous morphology with the scrapie-associated fibril²⁸, a candidate structure for the scrapie virus²⁹.

Standard curves, such as those in Fig. 4, are the preferred method of measurement in biophysical procedures such as sedimentation and electrophoresis where adequate standards exist. Now that D_{37} values can be calibrated against exact, sequence-based MWs for dozens of viruses³⁰, it no longer makes sense to attempt the estimation of nucleic acid MWs from first principles using the target theory. Moreover, past attempts to do so have seriously misled scrapie investigation.

ROBERT G. ROHWER

Department of Microbiology
and Immunology,
311 Lineberger Cancer Research
Center 237H,
University of North Carolina,
Chapel Hill,
North Carolina 27514, USA

1. Alper, T. *Nature* 317, 750 (1985).

2. Rohwer, R. G. *Nature* 308, 658–662 (1984).

3. Alper, T., Haig, D. A. & Clarke, M. C. *Biochem. biophys. Res. Commun.* 22, 278–284 (1966).

4. Rice, C. M. *et al. Science* 229, 726–733 (1985).

5. Fraenkel-Conrat, H. *Compreh. Virol.* 1, 18 (1974).

6. Bergold, G. J. & Weibel, J. *Virology* 17, 554–562 (1962).

7. Kaplan, H. S. & Moses, L. E. *Science* 145, 21–25 (1964).

8. McCrea, J. F. *Ann. N.Y. Acad. Sci.* 83, 692–705 (1960).

9. Wilson, D. E. *Radiat. Res.* 14, 796–802 (1961).

10. Syverton, J. T., Berry, G. P. & Warren, S. L. *J. exp. Med.* 74, 223–234 (1941).

11. Lea, D., Smith, K. M., Holmes, B. & Markham, R. *Parasitology* 36, 110–118 (1944).

12. Friedewald, W. F. & Anderson, R. S. *J. exp. Med.* 74, 463–487 (1941).

13. Woese, C. & Pollard, E. *Archs Biochem. Biophys.* 50, 354–367 (1954).

14. Hotz, G. *Molec. Biol. Biochem. Biophys.* 27, 304–311 (1978).

15. Ginoza, W. *Nature* 199, 453–456 (1963).

16. Ginoza, W. & Norman, A. *Nature* 179, 520–521 (1957).

17. Alper, T. & Haig, D. A. *J. gen. Virol.* 3, 157–166 (1968).

18. Gibbs, C. J. Jr, Gajdusek, D. C. & Latarjet, R. *Proc. natn. Acad. Sci. U.S.A.* 75, 6268–6270 (1978).

19. Terzi, M. *Nature* 191, 461–463 (1961).

20. Ginoza, W. *Meth. Virol.* 4, 139–209 (1968).

21. Schambra, F. E. & Hutchinson, F. *Radiat. Res.* 23, 514–526 (1964).

22. Tessman, I., Tessman, E. S. & Stent, G. S. *Virology* 4, 209–215 (1957).

23. Ginoza, W. & Miller, R. C. *Proc. natn. Acad. Sci. U.S.A.* 54, 551–558 (1965).

24. Hotz, G. & Muller, A. Z. *Naturforsch.* 16b, 282–283 (1961).

25. Latarjet, R., Muel, B., Haig, D. A., Clarke, M. C. & Alper, T. *Nature* 227, 1341–1343 (1970).

26. McLaren, A. D. *Photochem. Photobiol.* 8, 521–525 (1968).

27. Kleczkowski, A. & Govier, D. A. *Photochem. Photobiol.* 10, 53–59 (1969).

28. Merz, P. A., Somerville, R. A., Wisniewski, H. M. & Iqbal, K. *Acta neuropath.* 54, 63–74 (1981).

29. Merz, P. A. *et al. Science* 225, 437–440 (1984).

30. GenBank (Computer Systems Division, Bolt Beranek & Newman Inc., Cambridge, Massachusetts).

Applied Biological Sciences

Advances in Biotechnological Processes, Vol. 4. AVSHALOM MIZRAHI and ANTONIUS L. VAN WEZEL (eds). Alan R. Liss:1985. Pp.356. ISBN 0-8451-3203-2. £64.

Alcoholism and Substance Abuse: Strategies for Clinical Intervention. By THOMAS E. BRATTER and GARY G. FORREST. Collier Macmillan:1985. Pp.650. ISBN 0-02-904260-7. £40.

Annual Reports in Medicinal Chemistry, Vol. 20. D. M. BAILEY (ed.). Academic:1985. Pp.352. ISBN 0-12-040520-2. Np.

Antibodies: Protective, Destructive and Regulatory Role. F. MILGROM, C.J. ABEYOUNIS and B. ALBINI (eds). Karger:1985. Pp.462. ISBN 3-8055-3990-8. SwFr.375, DM 450, \$159.75.

Atlas of Positron Emission Tomography of the Brain. By W. D. HEISS, C. BELL, K. HERHOLTZ, G. PAWLICK, R. WAGNER and K. WIENHARD. Springer-Verlag:1985. (German and English.) Pp.130. Hbk ISBN 3-540-15636-4. DM 148.

"Autistic" Children: New Hope for a Cure. By N. and E.A. TINBERGEN. George Allen & Unwin:1985. Pp.362. Hbk ISBN 0-04-157010-3; pbk ISBN 0-04-157011-1. Pbk £9.95.

Carcinogenesis: A Comprehensive Survey. J. CARL BARRETT and RAYMOND W. TENNANT (eds). Raven:1985. Pp.461. ISBN 0-88167-116-9. \$70.50.

Communication Problems in Autism. ERIC SCHOPLER and GARY B. MESIBOV (eds). Plenum:1985. Pp.333. Hbk ISBN 0-306-41859-2. Np.

Comparison of Type I and Type II Diabetes: Similarities and Dissimilarities in Etiology, Pathogenesis, and Complications. MLADEN VRANIC, CHARLES H. HOLLENBERG and GEORGE STEINER (eds). Plenum:1985. Pp.349. ISBN 0-306-42066-X. Np.

Contemporary Topics in Immunobiology, Vol. 15: Immune Complexes and Human Cancer. FERNANDO A. SALINAS and MICHAEL G. HANNA Jr (eds). Plenum Press:1985. Pp.280. ISBN 0-306-41955-6. Np.

Decision-Making in General Practice. MICHAEL SHELDON, JOHN BROOKE and ALAN RECTOR (eds). Macmillan, London:1985. Pp.269. ISBN 0-333-36626-3. £25.

Dietary Fibre, Fibre-Depleted Foods and Disease. HUGH TROWELL, DENIS BURKITT and KENNETH HEATON (eds). Academic:1985. Pp.433. ISBN 0-12-701160-9. \$64.50, £49.50.

Directory of On-going Research in Cancer Epidemiology 1985. C.S. MUIR and G. WAGNER (eds). International Agency for Research on Cancer, World Health Organization:1985. Pp.756. Pbk ISBN 928-321169-3. £22.

Encounters with Qi: Exploring Chinese Medicine. By DAVID EINSENBURG. W.W. Norton:1985. Pp.254. ISBN 0-393-02213-7. \$16.95.

Ethical Issues in Preventative Medicine. SYPROS DOXIADIS (ed.). Martinus Nijhoff:1985. Pp.108. Hbk ISBN 90-247-3232-8. Np.

Evaluation of Large Ruminants for the Tropics. J.W. COPLAND (ed.). Australian Centre of International Agricultural Research (ACIAR), G.P.O. Box 1571, Canberra, Australia:1985. Pp.178. Pbk ISBN 0-949511-09-9. Np.

Experimental Research on Vinylidene Chloride Carcinogenesis. CESARE MALTONI et al. (eds). Princeton Scientific/Wiley:1985. Pp.229. ISBN 0-91131-38-8. £83.60.

Familial Cancer. HJ. MULLER, W. WEBER and T. KUTTAPA (eds). Karger:1985. Pp.292. Hbk ISBN 3-8055-4245-3. SwFr.185, DM 222, £57.90, \$78.75.

Free Radicals in Biology and Medicine. Life Chemistry Reports, Vol. 3, No. 1 and 2. J.V. BANNISTER, B. HALLIWELL and P. O'NEILL (eds). Harwood Academic Publishers:1985. Pbk ISBN 3-7186-0285-7. \$69.

Functional Mapping of the Brain in Vascular Disorders. WOLF-DIETER HEISS (ed.). Springer-Verlag:1985. Pp.126. Pbk ISBN 3-540-15801-4. DM 56.

Handbook of Clinical Behaviour Therapy with Adults. MICHEL HERSEN and ALAN S. BEL-LACK (eds). Plenum:1985. Pp.732. ISBN 0-306-41875-4. Np.

Handbook of Laboratory Health and Safety Measures. S.B. PAL (ed.). MTP Press:1985. Pp.391. ISBN 0-85200-766-3. £59.95.

Hepatitis B. ROBERT J. GERETY (ed.). Academic:1985. Pp.470. Hbk ISBN 0-12-280672-7. £68.50, \$79.

Interaction Theory in Forest Ecology and Management. By ROLFE A. LEARY. Martinus Nijhoff/Dr W. Junk:1985. Pp.219. Hbk ISBN 90-247-3220-4. Np.

Interferon 6, 1985. IAN GRESSER (ed.). Academic:1985. Pp.143. ISBN 0-12-302255-X. \$25, £16.

Interpretation of Endometrial Biopsies, 2nd Edn. By ANCEL BLAUSTEIN. Raven Press:1985. Pp.304. Hbk ISBN 0-88167-120-7. \$42.50.

Interpretation of Negative Epidemiological Evidence for Carcinogenicity. N.J. WALD and R. DOLL (eds). International Agency for Research on Cancer:1985. Pp.232. Hbk ISBN 92-832-1165-0. £20.

An Introduction to Human Ecology Research on Agricultural Systems in Southeast Asia. A. TERRY RAMBO and PERCY E. SAJISE (eds). University of the Philippines at Los Baños:1985. Pp.327. Pbk ISBN 0-86638-062-0. Np.

Iron Fortification of Foods. FERGUS M. CLYDESDALE and KATHRYN L. WIEMER (eds). Academic:1985. Pp.176. ISBN 0-12-177060-5. \$36.50, £32.50.

Laser Photobiology and Photomedicine. S. MARTELLUCCI and A.N. CHESTER (eds). Plenum:1985. Pp.334. Hbk ISBN 0-306-41915-7. Np.

Leukotrienes in Cardiovascular and Pulmonary Function. Progress in Clinical and Biological Research, Vol. 199. ALLAN M. LEFER and MARLYS H. GEE (eds). Alan R. Liss:1985. Pp.270. Hbk ISBN 0-8451-5049-9. £32.

The Meaning of Addiction: Compulsive Experience and Its Interpretation. By STANTON PEELE. Lexington Books:1985. Pp.203. ISBN 0-669-02952-1. £25.

Mechanisms of Resistance to Plant Diseases. R.S.S. FRASER (ed.). Martinus Nijhoff/Dr W. Junk Publishers:1985. Pp.462. Hbk ISBN 90-247-3204-2. Dfl.190, \$71.

The Medical and Biological Effects of Light. Annals of the New York Academy of Sciences, Vol. 453. RICHARD J. WURTMAN, MICHAEL L. BAUM and JOHN T. POTTS Jr (eds). The New York Academy of Sciences:1985. Pp.408. Hbk ISBN 0-89766-300-4; pbk ISBN 0-89766-301-2. Np.

Metal Ions in Biological Systems. Vol. 19. Antibiotics and Their Complexes. HELMUT SIGEL (ed.). Dekker:1985. Pp.429. ISBN 0-8247-7425-6. \$85.

Microsurgery of Cerebral Aneurysms: An Atlas. By ZENTARO ITO. Nishimura/Elsevier:1985. Pp.299. Hbk ISBN 90-219-3062-5. Dfl.600, \$222.25.

Modern Trends in Leukemia VI: New Results in Clinical and Biological Research Including Pediatric Oncology. ROLF NETH, ROBERT C. GALLO, MELVYN F. GREAVES and GRITTA JANKA (eds). Springer-Verlag:1985. Pp.521. Pbk ISBN 3-540-15329-2. DM 198.

Myelofibrosis: Pathophysiology and Clinical Management. S.M. LEWIS (ed.). Marcel Dekker:1985. Pp.207. ISBN 0-8247-7411-6. \$78.

New Cardiovascular Drugs, 1985. ALEXANDER SCRIBAN (ed.). Raven:1985. Pp.308. ISBN 0-88167-114-2. \$54.50.

Nuclear Medicine Annual 1985. LEONARD M. FREEMAN and HEIDI S. WEISSMANN (eds). Raven:1985. Pp.362. ISBN 0-88167-086-3. \$70.50.

The Patient Within the Patient: Problems in Perinatal Medicine. BRUCE K. YOUNG (ed.). Alan R. Liss:1985. Pp.240. Hbk ISBN 0-8451-1060-8. £30.

Pediatric Differential Diagnosis, 2nd Edn. By STEPHEN H. SHELDON and HOWARD B. LEVY. Raven Press:1985. Pp.209. Pbk ISBN 0-88167-112-6. Np.

Pharmaceuticals Among the Sunrise Industries: Proceedings of an Office of Health Economics Symposium. NICHOLAS WELLS (ed.). Croom Helm:1984. Pp.240. ISBN 0-7099-1947-6. £19.95.

Physical Chemistry: Principles and Applications in Biological Sciences, 2nd Edn. By IGNACIO TINOCO Jr., KENNETH SAUER and JAMES C. WANG. Prentice-Hall:1985. Pp.706. ISBN 0-13-666280-3. £37.40.

Practice and Theory of Enzyme Immunoassays. P. TIJSSEN (ed.). Elsevier:1985. Pp.549. Hbk ISBN 0-444-80634-2; pbk ISBN 0-444-80633-4. Hbk Dfl. 260, \$96.25; pbk Dfl. 85, \$31.50.

Principles of Metastasis. By L. WEISS. Academic:1985. Pp.425. ISBN 0-12-742820-8. \$52, £52.

Progress in Applied Microcirculation: Resistance Vessels: Physiology, Pharmacology and Hypertensive Pathology. M.J. MULVANY et al. (eds). Karger:1985. Pp.226. SwFr.140, DM 168, \$59.75.

Progress in Clinical and Biological Research, Vol. 178. Bluetongue and Related Orbiviruses. T. LYWOOD BARBER and MICHAEL M. JOCHIM (eds). Alan R. Liss:1985. Pp.772. ISBN 0-8451-5028-6. £84.

Progress in Clinical and Biological Research, Vol. 185A: EORTC Genitourinary Group, Monograph 2, Pt A. Therapeutic Principles in Metastatic Prostate Cancer. FRITZ H. SCHROEDER and BRIAN RICHARDS (eds). Alan R. Liss:1985. Pp.474. ISBN 0-8451-0188-9. £52.

Progress in Clinical and Biological Research, Vol. 185B: EORTC Genitourinary Group, Monograph 2, Pt B. Superficial Bladder Tumors. FRITZ H. SCHROEDER and BRIAN RICHARDS (eds). Alan R. Liss:1985. Pp.179. ISBN 0-8451-0189-7. £26.

The Pyrethroid Insecticides. J.P. LEAHEY (ed.). Taylor & Francis:1985. Pp.440. Hbk ISBN 0-85066-283-4. £45.

Radiolabeled Cellular Blood Elements: Pathophysiology, Techniques and Scintigraphic Applications. NATO ASI Series. M.L. THAKUR (ed.). Plenum:1985. Pp.434. Hbk ISBN 0-306-41935-1. Np.

Recent Achievements in Restorative Neurology, Vol. 1: Upper Motor Neuron Functions and Dysfunctions. JOHN ECCLES and MILAN R. DIMITRIJEVIC (eds). Karger:1985. Pp.345. ISBN 3-8055-4020-5. SwFr.279, DM 335, \$118.75, £87.20.

Surgical Pharmacology of the Eye. MARVIN L. SEARS and AHTI TARKKANEN (eds). Raven:1985. Pp.592. ISBN 0-88167-047-2. \$65.50.

The Tetracyclines. J.J. HLAVAKA and J.H. BOOTHE (eds). Springer-Verlag:1985. Pp.451. Hbk ISBN 3-540-15259-8. Np.

Toxic Susceptibility: Male/Female Differences. By EDWARD J. CALABRESE. Wiley:1985. Pp.336. ISBN 0-471-80903-9. £61.35.

Toxicology of Insecticides, 2nd Edn. By FUMIO MATSUMURA. Plenum:1985. Pp.598. Hbk ISBN 0-306-419-79-3. \$45.

Ultrasound Annual 1985. ROGERS C. SANDERS and MICHAEL C. HILL (eds). Raven Press:1985. Pp.381. Hbk ISBN 0-88167-143-6. \$76.50.

Your Healthy Heart. A Guide to How the Heart Works and How to Keep it Healthy. By CHRIST-IAAN BARNARD and PETER EVANS. Macdonald:1985. Pp.224. ISBN 0-356-10524-5. £12.95.

Psychology

Accounting for Aggression: Perspectives on Aggression and Violence. By GERDA SIANN. Allen & Unwin:1985. Pp.294. ISBN 0-04-301188-8. £6.95.

General

Analysis of Experiments with Missing Data. By YADOLAH DODGE. Wiley:1985. Pp.499. Hbk ISBN 0-471-88736-6. £40.85.

Aspects of Symmetry: Selected Erice Lectures. By SIDNEY COLEMAN. Cambridge University Press:1985. Pp.402. ISBN 0-521-26706-4. Np.

Constructing the Universe. By DAVID LAYZER. W.H. Freeman:1985. Pp.313. ISBN 0-7167-5003-1. \$27.95.

Das Phänomen des Polarlichts. By WILFRIED SCHRÖDER. Wissenschaftliche Buchgesellschaft, Darmstadt, FRG:1984. Pp.156. Pbk ISBN 3-534-08997-9. Pbk DM 37.50.

Digital Communications and Spread Spectrum Systems. By ROGER E. ZIEMER and ROGER L. PETERSON. Collier Macmillan:1985. Pp.750. Hbk ISBN 0-02-431670-9. £38.